

Normative and Pathological Personality Predictors of Generalized Conditioned Fear,
Instrumental Avoidance, and the Covariation of Generalized Fear and Avoidance

A DISSERTATION
SUBMITTED TO THE FACULTY OF
UNIVERSITY OF MINNESOTA
BY

Samuel Emerson Cooper

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Advisor: Shmuel Lissek, Ph.D.

August 2019

Acknowledgements

First and foremost, thank you to Dr. Shmuel Lissek, my mentor and principal advocate over the last six years. Shmuel, I appreciate the many hours we have spent discussing conditioning, anxiety, experimental techniques, and all the rest. I look forward to many more hours digging into the intricacies of these and other esoterica in the future. I also want to acknowledge your flexibility, willingness to dive into new literatures related to my interests, and for caring about me, my values, and my goals beyond the research lab. That said, we will forever need to agree to disagree on the question of whether New York bagels are superior to others (they are, and it's not even a question).

I also want to thank the many research mentors, advisors, and professors who have provided training, skills, guidance, and food-for-thought over the years. I carry lessons from all of you, and I look for opportunities to pay it forward whenever I can. Thank you for providing the opportunity to participate in this wonderful field and the foundation from which grow in its service.

Broadly, I'd like to thank my colleagues and friends throughout my graduate career; I appreciate you all. I want to offer specific acknowledgment and thanks to the following graduate student colleagues: Hannah Berg, for getting this project to the finish line and offering new perspectives on the paradigm; John Gaffney, for getting me hooked on personality science and co-piloting this study with me; Abbey Hammell, for headache-inducing (but immensely beneficial) stats conversations; and Chris Hunt, for jumpstarting the analyses of these data and many conversations regarding complex interpretations. Thank you for always allotting 5 (to 10) minutes for me to make dumb jokes or reference obscure sports figures before our meetings turned to things that actually matter. I also want to acknowledge colleagues from the last leg of my graduate journey, Dr. Dan Lakin and Dr. Matt Barstead, for your extremely helpful statistical advice and guidance, as well as general encouragement, during the home stretch.

Many research assistants and coordinators contributed to this dissertation study, from conducting many (many) testing sessions, to data entry and management, to training and

protocol writing, and all the other things that are vital to this type of project. Everyone's contribution was invaluable. I also want to specifically thank Melissa Hartnell and Jack Ross for their efforts that were the very definition of "above and beyond" – even if we never did solve the mystery of "ZQ".

It's possible that we (I) don't take the time to sufficiently acknowledge and voice appreciation for what we (I) value. Therefore, I would like to add some less common acknowledgements. Thank you to my constant companions during the writing process, the Wu Tang Clan. Thank you to the judo community. Thank you to the many dogs in my life, past and present, both those that I am on a first name basis with and those with whom I exchange pleasantries but can't really say I know all that well. Thank you to Wally's and Maiwand Grill, and for 100% buying my explanation that the order of extra bread is for a friend. And finally, I would like to acknowledge the many fine coffee shops and their staff in Minneapolis, Baltimore, New York City, and Detroit, that served as my makeshift offices, with a special shout out to both Butter of Minneapolis and Baby's on Fire of Baltimore. I am sorry for taking up two tables most of the time.

I also want to thank my family, as well as my friends who predate graduate school. Your support throughout this period in my life is what kept me going. I would also like to specifically thank my parents, Judith Wineman, and Melissa Brophy and Matthew Cooper, because you were always there for me, taught me everything, and if I didn't add in this specific "thank you" I would never hear the end of it.

Finally, I value my clinical scientist identity and believe it has improved me as a researcher and as a person. I also cannot separate my clinical identity from my training at the Minneapolis and Baltimore VA medical centers, most notably as a practicum student and then as an intern within PTSD and trauma clinics. Thus, thank you to my VA clinical supervisors for your training, guidance, wisdom, and for helping me keep things in perspective. And in closing, my sincere gratitude and thanks go out to those veterans with whom I had the privilege to work: you were all inspirations and I will endeavor to conduct research that makes a difference, for veterans and civilians alike.

Dedication

For all of the grandparents, and for Peaches.

Abstract

Generalization of classically conditioned fear to novel stimuli that closely resemble a danger cue (CS+) can be adaptive, as the brief mobilization of biological defense reactions is typically not harmful to the organism and the outcome can, in some contexts, mean the difference between life and death. Generalization becomes maladaptive when fear is generalized to 1) to stimuli with an established safe signal value (CS-) and 2) to an excessively wide range of benign stimuli that inconsequentially resemble the danger cue (generalization stimuli [GS]) as excessive defense mobilization eventually becomes harmful to the organism. Mechanistic conditioning models of human anxiety pathology have termed this maladaptive form of generalization as *overgeneralization*, and experimental studies have established overgeneralization as a correlate of clinical anxiety (e.g., the anxiety disorders). These models have also, until recently, largely discounted the pathological contribution of instrumental avoidance of feared stimuli. This is in stark contrast to clinical models of anxiety pathology, which establish that the most severe forms of anxiety disorder involve excessive avoidance that results in loss of valued activity and opportunity to extinguish fear, and links this avoidance to individual differences in a variety of personality traits. Recent mechanistic work has partially addressed this gap and investigated the relationship between generalized fear and generalized avoidance, but has largely not incorporated individual difference variables. The current investigation furthers the merging of mechanistic conditioning and clinical models in this area by testing how broadband individual differences (e.g., personality traits) ranging from normative to pathological can improve prediction of instrumental avoidance from generalized fear. Candidate personality variables include those related to Conscientiousness and Extraversion, both traits that are linked to learning and approach systems. The method for this investigation involved lab-based assessment using established conditioning paradigms with behavioral and psychophysiological indicators, as well as multidimensional self-report inventories and a multilevel modeling analytic approach to facilitate more precise testing of personality-related hypotheses. Results indicate that 1) multiple measures of pathological negative affect are related to increased fear generalization and facilitate a maladaptive fear-avoidance relations; 2) Extraversion-related variables generally buffer against fear-avoidance covariation, whereas

pathologically low Extraversion (detachment) facilitates the fear-avoidance relation; 3) Conscientiousness-related variables both facilitate and inhibit the fear-avoidance relation, depending on context; and 4) the relationship between the personality variables, generalized fear, and avoidance depends partially on how the fear metric is operationalized (e.g., physiologically or behaviorally). These results are discussed within a framework of improving methodology for investigations that combine conditioning and individual differences approaches and using this type of work to inform translational efforts to further refine and personalize treatments for anxiety and trauma-related psychopathology.

Table of Contents

<i>List of Tables</i>	<i>viii</i>
<i>List of Figures</i>	<i>ix</i>
 <i>Introduction.....</i>	 <i>1</i>
Pavlovian Fear Conditioning and Generalization of Fear	3
Instrumental Avoidance and Approach-Avoidance Conflict	8
Covariation of Generalized Conditioned Fear and Instrumental Avoidance.	11
APIC-G and APIC-CS+	13
Normative and Pathological Individual Differences in Generalized Fear,	
Generalized Avoidance, and APIC-G/APIC-CS+.....	19
The importance of inter-individual personality differences.....	20
Review of individual differences work in generalization and related studies.	22
Methodology of Individual Differences	67
Conceptual rationale for a dimensional individual differences approach.....	67
Statistical rationale for a dimensional individual differences approach.	69
Summary and Conclusions.....	77
The Current Study and Specific Aims	84
Specific aims.	85
Hypotheses and exploratory testing approach.	86
<i>Method.....</i>	<i>97</i>
Participants.....	97
Materials	99
Trait-level questionnaires.....	99
State level questionnaires.....	105
Physiological apparatus	106
Pavlovian-Instrumental Generalization (PIG) task.	106
Procedure.....	113
Data Processing and Statistical Analysis Plan.....	114
Startle data reduction and preparation.	114
Self-report data reduction and preparation.	114
Preliminary statistical analyses.	115
Main statistical analyses.	116
<i>Results.....</i>	<i>132</i>
Preliminary Analyses.....	132
Manipulation checks.	132
Participant characteristics.	136
Main analyses	141

Aim 1 (A1): Identifying personality predictors of generalization.	141
A1: Startle.	141
A1: Risk ratings.	147
A1: Avoidance.	156
Interim summary for Aim 1.	169
Aim 2 (A2): Testing improved prediction of avoidance in APIC models.	173
A2: Startle predicting avoidance.	174
A2: Risk ratings predicting avoidance.	176
Interim summary for Aim 2.	177
Aim 3 (A3): Moderation of APIC by dispositional traits.	179
A3: Startle predicting avoidance.	180
A3: Risk ratings predicting avoidance.	182
Interim summary for Aim 3.	189
Aim 4 (A4): Specific APIC-G and APIC-CS+ effects.	191
A4: APIC-G: Startle predicting avoidance.	192
A4: APIC-G: Risk ratings predicting avoidance.	192
A4: APIC-CS+: Startle predicting avoidance.	202
A4: APIC-CS+: Risk ratings predicting avoidance.	202
Interim summary for Aim 4.	202
Discussion.	205
Aim 1: Identifying personality predictors of generalization.	205
Aim 2: Testing improved prediction of avoidance in APIC models.	219
Aims 3 and 4: Moderators of APIC, APIC-G, and APIC-CS+.	224
Overall Synthesis of Results.	237
Broadband and narrowband negative affect traits.	237
Conscientiousness and related variables.	239
Extraversion and related variables.	242
Across-personality trends.	244
Benefits of a multi-trait, multi-level approach.	246
Limitations.	247
Experimental limitations.	247
Statistical limitations.	252
Sample and individual difference-related limitations.	257
Future Directions.	260
Optimizing the PIG.	260
Incorporating additional measures.	264
Implications for Clinical Science and Intervention.	266
References.	273
Appendix A.	319
Appendix B.	320

List of Tables

Table 1. Participant demographics.....	98
Table 2. Exclusion conditions and totals	99
Table 3. Descriptive statistics for broadband and narrowband measures.	136
Table 4. Trait variable correlations	139
Table 5. Task variable correlations.....	140
Table 6. Summary table for Aim 1: Significant Main Effects and Interactions by Personality Variable.....	172
Table 7. Summary table for Aim 2: Significant Improvements in APIC Model Fit and Corresponding Effect Size (Marginal R^2).....	178
Table 8. Summary table for Aim 3: Significant APIC Moderators	190
Table 9. Summary table for Aim 4: Significant APIC-G Moderators.....	204

List of Figures

Figure 1. Example of generalization gradients.	8
Figure 2. Overview of Pavlovian-Instrumental Generalization (PIG) paradigm.....	108
Figure 3. Example trials from the acquisition phase and generalization phase (both Pavlovian and Instrumental trials) of the PIG paradigm..	111
Figure 4. Acquisition manipulation check results (rmANOVAs).....	133
Figure 5. Generalization manipulation check results (rmANOVAs).....	135
Figure 6. Plotted actual and fitted startle values for Extraversion x Stimulus ² (quadratic) interaction (generalization MLM).....	143
Figure 7. Plotted actual and fitted startle values for Assertiveness x Stimulus ² (quadratic) interaction (generalization MLM).....	145
Figure 8. Plotted actual and fitted startle values for TF-44 x Stimulus ² (quadratic) interaction (generalization MLM).....	146
Figure 9. Plotted actual and fitted startle values for IUSF x Stimulus ² (quadratic) interaction (generalization MLM).....	147
Figure 10. Plotted actual and fitted risk rating values for Neuroticism x Stimulus interaction (generalization MLM).....	148
Figure 11. Plotted actual and fitted risk rating values for Negative Affectivity x Stimulus interaction (generalization MLM).....	151
Figure 12. Plotted actual and fitted risk rating values for Detachment x Stimulus interaction (generalization MLM).....	152
Figure 13. Plotted actual and fitted risk rating values for Distress-PB x Stimulus interaction (generalization MLM).....	154
Figure 14. Plotted actual and fitted risk rating values for Enthusiasm-PB x Stimulus interaction (generalization MLM).....	155
Figure 15. Plotted actual and fitted risk rating values for IUSF x Stimulus ² (quadratic) interaction (generalization MLM).....	156
Figure 16. Plotted actual and fitted avoidance % for Neuroticism x Stimulus ² (quadratic) interaction (generalization MLM).....	157
Figure 17. Plotted actual and fitted avoidance % for Volatility x Stimulus ² (quadratic) interaction (generalization MLM).....	158
Figure 18. Plotted actual and fitted avoidance % for Conscientiousness x Stimulus ² (quadratic) interaction (generalization MLM).....	159
Figure 19. Plotted actual and fitted avoidance % for Orderliness x Stimulus ² (quadratic) interaction (generalization MLM).....	160
Figure 20. Plotted actual and fitted avoidance % for Negative Affectivity x Stimulus ² (quadratic) interaction (generalization MLM).....	162
Figure 21. Plotted actual and fitted avoidance % for Detachment x Stimulus ² (quadratic) interaction (generalization MLM).....	164
Figure 22. Plotted actual and fitted avoidance % for Distress-PB x Stimulus ² (quadratic) interaction (generalization MLM).....	165
Figure 23. Plotted actual and fitted avoidance % for Orderliness-PB x Stimulus ² (quadratic) interaction (generalization MLM).....	167
Figure 24. Plotted actual and fitted avoidance % for ASI x Stimulus ² (quadratic) interaction (generalization MLM).....	169

Figure 25. Plotted actual and fitted avoidance % for STAI-T x Stimulus ² (quadratic) interaction (generalization MLM).....	169
Figure 26. Scatterplots and simple slopes for Orderliness x Stimulus x Startle interaction (APIC MLM)	169
Figure 27. Scatterplots and simple slopes for Conscientiousness x Stimulus x Risk Rating interaction (APIC MLM).	183
Figure 28. Scatterplots and simple slopes for Orderliness x Stimulus x Risk Rating interaction (APIC MLM).	184
Figure 29. Scatterplots and simple slopes for Extraversion x Stimulus x Risk Rating interaction (APIC MLM).	186
Figure 30. Scatterplots and simple slopes for Assertiveness x Stimulus x Risk Rating interaction (APIC MLM).	188
Figure 31. Scatterplots and simple slopes for Extraversion x Stimulus x Risk Rating interaction (APIC-G MLM).	193
Figure 32. Scatterplots and simple slopes for Enthusiasm x Stimulus x Risk Rating interaction (APIC-G MLM).	195
Figure 33. Scatterplots and simple slopes for Negative Affectivity x Stimulus x Risk Rating interaction (APIC-G MLM).	196
Figure 34. Scatterplots and simple slopes for Detachment x Stimulus x Risk Rating interaction (APIC-G MLM).	198
Figure 35. Scatterplots and simple slopes for Assertiveness-PB x Stimulus x Risk Rating interaction (APIC-G MLM).	199
Figure 36. Scatterplots and simple slopes for Enthusiasm-PB x Stimulus x Risk Rating interaction (APIC-G MLM).	201

Introduction

Conditioned fear and avoidance models of anxiety pathology have led to valuable advances that have directly contributed to diagnostic frameworks and effective interventions that meaningfully alleviate symptoms, most notably exposure therapy and other components of the cognitive behavioral framework (Bouton & Swartzentruber, 1991; Mineka & Zinbarg, 2006; Rachman, 2015). These advances are partially attributable to sustained empirical focus, increased experimental precision, and advancements in behavioral and biological measurement in the conditioning field over the last century (Craske, Hermans, & Vervliet, 2018; Hermans, Craske, Mineka, & Lovibond, 2006). Despite this, as a field we can do better: a significant minority of people with debilitating anxiety and trauma-related disorders do not respond to empirically-supported treatment (Arch & Craske, 2009; Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014; McNally, 2007; Rothbaum, Meadows, Resick, & Foy, 2000; Steenkamp, Litz, Hoge, & Marmar, 2015) and precision-medicine is still a distal goal for clinical psychology and psychiatry (Cuijpers, Ebert, Acarturk, Andersson, & Cristea, 2016; Ozomaro, Wahlestedt, & Nemeroff, 2013). It is still difficult to know who will complete or positively respond to a “gold-standard” psychotherapy and what factors will positively or negatively impact their course of treatment (Gutner, Gallagher, Baker, Sloan, & Resick, 2015; Ong, Lee, & Twohig, 2018; Schneider, Arch, & Wolitzky-Taylor, 2015).

From the clinician point of view, this is fairly understandable: many treatments are developed using carefully curated samples comprised of groups of “modal participants” that likely do not reflect the heterogeneity of the clinic population (Barber,

2009; Najavits, 2015). Consider prolonged exposure (PE), a gold-standard treatment for posttraumatic stress disorder (PTSD) with strong roots in conditioning theory and research which inform PE's focus on fear and avoidance reduction through repeated imaginal and behavioral exposure (Foa & Kozak, 1986; Foa, Hembree, & Rothbaum, 2007). However, the conditioning framework that underlies PE gives no indication of how to handle inter-individual differences in key non-fear related traits and states that can have a profound impact on treatment adherence and outcome, such as strong levels of impulsivity, anhedonia, or disorganization (Minnen, Harned, Zoellner, & Mills, 2012; Reger et al., 2013). Although there are clearly developments and techniques from outside the tradition of conditioning research that clinicians use to handle these issues while maintaining treatment fidelity (e.g., Arkowitz & Westra, 2004; Hundt, Barrera, Arney, & Stanley, 2017), there still remains the concern that fear conditioning investigations continue to be limited in their applicability to "real world" heterogeneity of individual differences seen in patient populations (Beckers, Krypotos, Boddez, Effting, & Kindt, 2013; Lonsdorf & Merz, 2017). Accordingly, for the fear conditioning field to continue to have a meaningful impact on clinical research and outcomes, the issue of "real world" heterogeneity of individual differences, and how they relate to conditioned fear and avoidance processes, requires redress.

This dissertation represents a step towards this goal. We will first review the conditioned fear and avoidance literature, with a focus on generalization of fear and avoidance, which has been posited as a core pathogenic mechanism of anxiety and trauma pathology (Dymond, Dunsmoor, Vervliet, Roche, & Hermans, 2014; LeDoux, Moscarello, Sears, & Campese, 2017; Lissek, 2012; Pittig, Treanor, LeBeau, & Craske,

2018). Next, we discuss the association and potential covariation between generalized fear and avoidance, and establish the relative lack of empirical investigations of individual differences in this relationship, despite its importance to understanding and treating anxiety and trauma pathology. We then review the available evidence for individual differences in fear generalization, avoidance, and their covariation, with a focus on personality variables (using the general definition of a relatively stable disposition towards a behavioral outcome, e.g., Tellegen, 1991) and clinical disorders (e.g., DSM disorders; American Psychiatric Association, 2013). The bulk of the dissertation then describes the first large-scale study of personality differences in generalized fear, avoidance, and their covariation. We end with a discussion of the study results, contextualize them in both the conditioning and clinical research traditions, and conclude with implications for future research and clinical endeavors.

Pavlovian Fear Conditioning and Generalization of Fear

Pavlovian fear conditioning, also referred to as classical fear conditioning, is an associative learning process in which an aversive unconditioned stimulus (US) is paired with a previously benign cue that now becomes a conditioned danger cue (CS+), with the CS+ eliciting fear even when not paired with the US (Pavlov, 1927). This form of fear learning is highly adaptive in many contexts, as it promotes proper mobilization of defensive responses to threats in the environment (e.g., Bradley, Moulder, & Lang, 2005; Lang, McTeague, & Bradley, 2014). For example, it is adaptive for a child to be afraid of

a large Rottweiler¹ dog after that dog lunges and bites him² (with the bite as the US and the Rottweiler becoming the CS+) and continues to experience fear when encountering that specific dog. The child has quickly conditioned to associate a fear response with this stimulus (i.e., automatically experiences fear when encountering the CS+), and this fear response serves as an aversive signal to inform future behavior. Given the context, this fear is adaptive because the Rottweiler is a genuine threat. Another important element of fear conditioning is that some stimuli in the environment are never paired with the US, and become conditioned safety cues (CS-), as they signal safety or absence of threat. Continuing with the previous example, the child might also encounter a different dog that is dissimilar to the Rottweiler, such as a toy poodle. The poodle never attempts to bite the child, and the child does not develop a fear response to the toy poodle. The lack of fear response is adaptive in this context, as the poodle is safe and a signal for further defensive responding is not needed. Discriminating between danger and safety cues, such as the ones described, is crucial for survival across species and is also key for optimizing internal reactions and external behavior to maximize resource allocation and usage (Maren, 2001; Rachman, 1991). It also serves as the foundation for basic experimental investigations of Pavlovian fear conditioning, referred to as discrimination conditioning paradigms. This paradigm has served as the basic framework for investigating different fear conditioning processes and mechanisms, with a large number of conditioning

¹ The author would like to note that he does not host any animus towards Rottweilers – they are simply a well-known dog breed with an easily recognizable name and perceptual features that lend themselves well to examples of fear generalization.

² A gendered pronoun is used to refer to the child in this example and examples that follow to streamline prose in this dissertation; the implied gender is not meaningful or relevant within these examples and a coin flip was used to select between male or female pronouns.

paradigms using it as a foundation for more complex experimental manipulations (Lonsdorf et al., 2017).

One mechanism of particular empirical interest is generalization of Pavlovian-conditioned fear (Kalish, 1969; Mackintosh, 1974; J. B. Watson & Rayner, 1920), in which fear is elicited by stimuli that perceptually or conceptually resemble the CS+ but are not dangerous (i.e., generalization stimuli, or GS). The importance of studying fear generalization stems from both its implication as a mechanism underlying anxiety disorders and its relevance to optimizing behavior in a world that contains many ambiguous stimuli with differing signal values (Dymond, Dunsmoor, et al., 2014; Lissek et al., 2005). To continue the Rottweiler example, the child's fear of the Rottweiler might generalize to dog breeds that are very similar in coloring and shape, despite some key differences in physical attributes (e.g., Dobermans, which have the same coloring and are of similar height and length, but are generally less muscular and have a slimmer facial structure than Rottweilers). Fear generalization can also be viewed from the lens of adaptive vs. maladaptive depending on the context. For example, it might also be adaptive to have some fear to unfamiliar Rottweilers that show warning signs of danger (e.g., exposed teeth, growling) but have not actually attempted an attack (i.e., a stimulus that is very similar to the CS+ but has not yet been paired with a US), as the brief mobilization of biological defense reactions is typically not harmful and the outcome can, in some contexts, mean the difference between serious harm or death and remaining safe. Conversely, it is likely maladaptive for a child to respond fearfully to a Doberman that has never displayed signs of threat or aggression (i.e., a GS), as there is no evidence that the dog is actually dangerous, and the fear response is unnecessarily activated in a safe

context. Further, the child might excessively generalize his fear and continue to respond fearfully to non-dangerous dogs that only superficially resemble a Rottweiler (e.g., black coloring, similar body profile) or even a Doberman, and have many features that are quite distinct from a Rottweiler (e.g. different face profile, much smaller size). This degree of increased fear generalization, deemed *overgeneralization*, can be conceptualized as a particularly maladaptive instantiation of the fear generalization process (Dunsmoor, Kroes, Braren, & Phelps, 2017; Dunsmoor & Paz, 2015; Lissek, 2012).

Fear generalization has increasingly been studied using laboratory-based paradigms (Dymond, Dunsmoor, et al., 2014; Lissek, Biggs, et al., 2008; Vervliet & Geens, 2014; Vervliet, Vansteenwegen, & Eelen, 2004). In fear generalization paradigms, participants complete an initial conditioning phase (typically referred to as the acquisition or conditioning phase) during which the signal-value of a CS+ and CS- are learned. The CSs are typically simple visual stimuli that are clearly distinguishable from each other (e.g., a large and a small circle)³, and the US is typically a simple aversive stimulus, such as a mild shock. In the next phase (typically referred to as the generalization phase), GSs are presented to the participant. Exact procedural and stimulus characteristics vary by study (e.g., US reinforcement rate, number of GS presentations and permutations), but a key unifying aspect of these paradigms is that the GSs will parametrically vary along a continuum of similarity that is anchored by the CS+ and CS- (see Figure 1 for example

³ It should be noted that a growing number of studies are investigating conceptual, as opposed to perceptual, fear generalization. In these studies, stimuli differ in terms of their conceptual similarity (e.g., birds and bats), as opposed to their perceptual similarity. Although beyond the scope of this dissertation, given both that the current study is one of perceptual generalization and that field has not yet begun to study conceptual generalization in association with avoidance, readers are referred to Dunsmoor & Murphy (2015) for a review on the importance of continued study of conceptual generalization.

from Lissek et al., 2014). This experimental manipulation allows fear generalization to be operationalized by a response gradient (i.e., line representing response magnitude at each stimulus) from the CS+, to the GSs, and then to the CS-. More precipitous declines in response magnitude from the CS+ to the CS- represent less generalization, and shallower response gradients that adhere closer to a perfect linear decline represent more generalization. These paradigm properties also allow for testing of overgeneralization, as this can be conceptualized as a response gradient with GS slope components that exceed the response expected as part of a perfect linear decline – in other words, overgeneralization is seen when one or more GS responses are elevated above a hypothetical straight line that connects the CS+ and CS- (see Figure 1 for graphical representation). Those with gradients representing higher levels of generalization can be conceptualized as experiencing maladaptive fear, as the fear response is activated unnecessarily due to the lack of danger represented by the presented stimulus and, in some cases, can lead to chronically activated fear and stress states associated with a range of negative emotional and medical health outcomes (e.g., Besnard & Sahay, 2016a, 2016b; Kudielka & Kirschbaum, 2005).

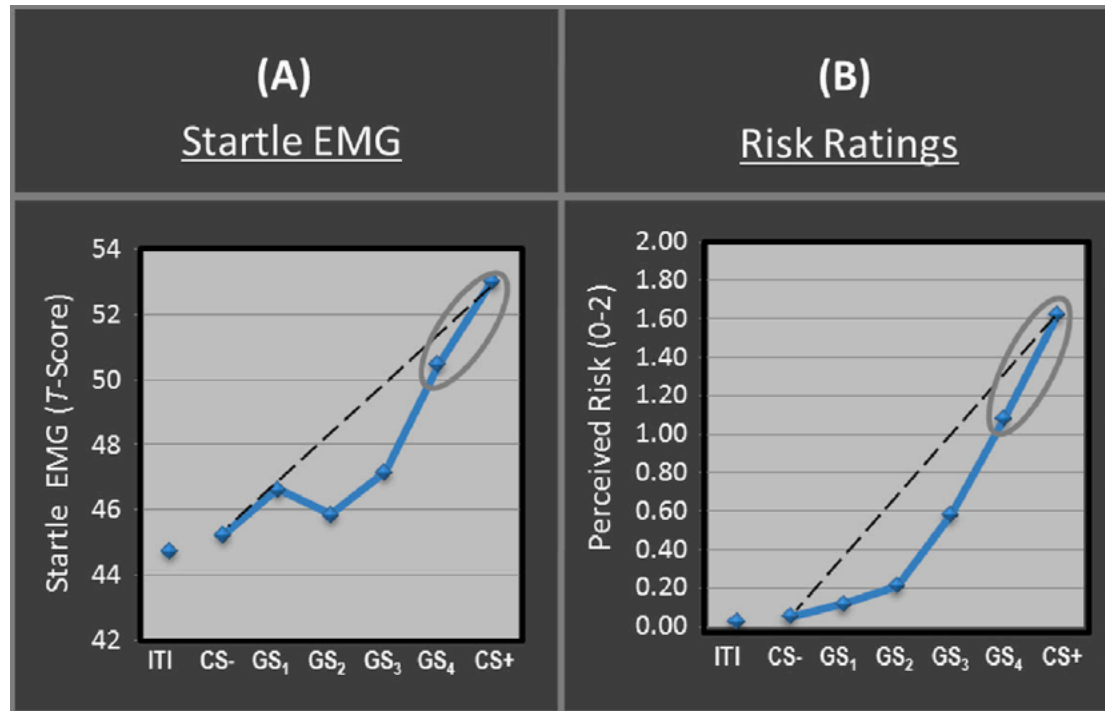


Figure 1. Example of generalization gradients from participants without psychopathology (i.e., “healthy control participants”). Physiological (startle electromyography) and behavioral (risk rating) gradients are presented. Gradient images adapted from a scientific presentation given by Dr. Lissek, with his permission. CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimuli; ITI = inter-trial interval; RR = risk rating.

Instrumental Avoidance and Approach-Avoidance Conflict

Although it is possible to conceptualize the maladaptive consequences of excessive fear generalization solely through internal consequences, such as the aforementioned health outcomes, it leads to an incomplete picture. In particular, it is missing something that is quite clear and intuitive if we continue the example of the Rottweiler and child: if the child continues to maintain fear to both the Rottweiler and dogs that are similar, he will likely seek to avoid those dogs. To review this topic, we temporarily step away from the Pavlovian fear conditioning literature and laboratory paradigms and consider avoidance as a basic psychological construct. Avoidance refers to the mobilization of resources (e.g., physical energy, time spent planning) in the service of

behavior that reduces the organisms contact with or moves it away from negatively-valenced stimuli or outcomes (Corr, 2013; Elliot, 2006). Avoidance of aversive stimuli in the environment can be accomplished either through active avoidance (e.g., when the child sees a feared dog down the street he walks the other way) or passive avoidance (e.g., deciding not to leave the house if the weather is fair and dogs might be outside walking), with both forms serving the goal of reducing fear and perceived chance of harm (for reviews, see Arnaudova, Kindt, Fanselow, & Beckers, 2017; Krypotos, 2015; Pittig et al., 2018). Both these forms of avoidance are variations on the basic conditioning principle of instrumental (or operant) conditioning, in which a behavioral outcome is shaped through prior experience, either positive or negative (Skinner, 1937, 1963)⁴. In the case of aversive stimuli, instrumental avoidance is conceptualized as a negative reinforcement process in which escape or reduction of an aversive consequence promotes future avoidance (e.g., Fernando, Urcelay, Mar, Dickinson, & Robbins, 2014; Mkrtchian, Aylward, Dayan, Roiser, & Robinson, 2017; Sidman, 1962).

Just as with Pavlovian fear conditioning, we can consider instrumental avoidance of aversive stimuli as adaptive or maladaptive for the organism (Lommen et al., 2017; van Meurs, Wiggert, Wicker, & Lissek, 2014). In both instances of the avoidance described in the ongoing example, the determination of adaptive vs maladaptive outcome is dependent on the specific context of the avoidance: if the child actively avoids when the dangerous Rottweiler (CS+) is present, it is an adaptive decision and harm is avoided;

⁴ The current dissertation focuses on active avoidance, both because it is tested in the current study and because the majority of prior human clinical work has focused on paradigms utilizing active avoidance (for review, see Arnaudova, Kindt, Fanselow, & Beckers, 2017).

if the child avoids a dissimilar looking and benign toy poodle (CS-) then, at first glance, it is maladaptive and avoidance was unnecessary due to the lack of danger.

However, simply terming these outcomes as adaptive or maladaptive based on the dichotomy of “was avoidance necessary or not” is problematic. First, it ignores that from the child’s view, both instances of avoidance resulted in a favorable outcome for him: his fear is likely reduced or has not increased, and the child would view both instances of avoidance as adaptive. To fully appreciate the relevant adaptive or maladaptive consequences in this example, especially as they pertain to future learning for the child, we need to consider motivations other than the motivation to avoid harm and reduce fear, and if there is a conflict created by competing motivations (Corr, 2013; Pittig & Dehler, 2018; Pittig, Treanor, et al., 2018). The most commonly studied motivational conflict across species is the risk-reward or cost-benefit conflict, which hinges on the conflict experienced when attempting to maximize gains while minimizing losses and is well-documented on both the behavioral and neural levels (for reviews, see Levy & Glimcher, 2012; Reyna & Huettel, 2014). This conflict can emerge when the decision hinges on outcomes of the same valence or of opposing valences (Aupperle & Paulus, 2010; Gray, 1990), and is driven by the challenge of either deciding between two different rewards (e.g., choosing between ice cream or cake), deciding between two different costs/risks (e.g., running on the treadmill or going for a swim to lose weight), or from the deciding if a reward outweighs the negative consequence (e.g., deciding whether to eat ice cream when trying to lose weight). A specific form of the latter type of conflict that is the most relevant to the ongoing example and the current study is the approach-avoidance conflict. Proposed as a core structure of human motivation (Elliot, 2006; Elliot & Thrash, 2002),

the approach-avoidance conflict results when an approach action is associated with both potential reward and potential risk or harm (e.g., Aupperle, Sullivan, Melrose, Paulus, & Stein, 2011; Aupperle & Paulus, 2010; Carver & White, 1994; Elliot, 2006; Gray, 1990). Resolution of the conflict results from either a decision to approach or a decision to avoid; it is not contingent on actually receiving the reward or punishment. This provides context for a complete definition of maladaptive avoidance: if avoidance is enacted at the expense of reward when there is no risk of harm, it is maladaptive. The opposite situation can also serve as a definition of adaptive avoidance (avoiding when there is genuine risk and forgoing reward). In both cases, there is an assumption that the risk and reward are sufficiently hedonically salient to motivate an actual approach-avoidance conflict (Aupperle & Paulus, 2010).

Bringing us back to the ongoing example, we see that adding a reward or approach component helps determine the adaptiveness of the child's dog-related decisions. For example, if the child is on his way to school and sees the Rottweiler (CS+) walking down the street he is on, and he then decides to take another street that is out of his way and causes him to be late, this might be considered adaptive: it is likely better to arrive at school slightly late than receive a painful dog bite. However, if the dog the child encounters is the toy poodle (CS-) and he takes the other road and arrives late to school than this outcome can be deemed as maladaptive, as avoidance resulted in an unnecessary loss of reward (for this example, the child is very studious and values arriving at school on time and engaging in education).

Covariation of Generalized Conditioned Fear and Instrumental Avoidance.

Returning to a conditioning framework, we can conceptualize how Pavlovian conditioning and instrumental avoidance are two interconnected parts of how fear is learned and maintained: CS+s are signals for avoidance due to their negative valence, and the absence of fear as a result of avoiding a CS+ is a negative reinforcer for future avoidance. This view was first widely posited by O. H. Mowrer and termed two-factor learning theory (Mowrer, 1951). In two-factor learning theory, Mowrer proposes that 1) initially benign stimuli acquire threat valence through Pavlovian conditioning and then 2) the reduction of fear accomplished through avoidance of the conditioned stimulus serves as the “reward of avoidance” (potentially experienced as relief, for a seminal review on this distinction see Hull, 1935) and reinforces future avoidance (hence, instrumental avoidance). Although this theory was met with considerable criticism and offering of contradictory empirical evidence that led to a temporary but lengthy reduction in the use of avoidance paradigms in the conditioning field (for a narrative history and summary of critique, see both LeDoux et al., 2017 and Rachman, 2015) and alternative theories emerged to fill in perceived gaps in its logic or as advances in its basic tenets (e.g., expectancy/cognitive models and meaning-making models, e.g., Lovibond, Saunders, Weidemann, & Mitchell, 2008; Seligman & Johnston, 1973; Thompson, 1981), the basic two-factor structure was maintained by some theorists as a parsimonious and fundamentally sound, if not overly precise, description of one of the commonly observed relationships between fear and avoidance (Allen, Handy, Miller, & Servatius, 2019; Lissek & van Meurs, 2014; Rescorla & Solomon, 1967; Sidman, 1953; Solomon & Wynne, 1954). For the current endeavor, we view it as a helpful framework for understanding possible determinants of psychopathological anxiety and fear responding,

but default to a simpler assertion as to how generalization of Pavlovian fear can lead to generalization of instrumental avoidance: as conditioned fear towards a GS increases, so does the likelihood of avoidance. In other words, as fear varies, so does avoidance, and their relationship can be described as one of covariation, herein referred to as Aversive Pavlovian-Instrumental Covariation (APIC), which is subdivided into APIC-CS+ (covariation of Pavlovian and instrumental responses to the danger cue) and, most importantly for this dissertation, into APIC-G (covariation of Pavlovian and instrumental responses to generalization stimuli). Below, we expand on the relevance of APIC-G and its appearances in the empirical literature.

APIC-G and APIC-CS+

Empirical documentation of APIC-G is still relatively sparse. This is somewhat surprising, given that APIC-G has been proposed in the experimental literature as a primary pathogenic mechanism that converts the distress of generalized fear into impairment related to maladaptive avoidance, and might explain why overgeneralized fear is found in the majority of cited case-control anxiety disorder studies. To date, only four⁵ studies have experimentally tested APIC-G using generalization paradigms employing stimuli on a continuum of similarity (Arnaudova, Kryptos, Effting, Kindt, & Beckers, 2017; Boyle, Roche, Dymond, & Hermans, 2016; Cameron, Schlund, & Dymond, 2015; van Meurs et al., 2014). Shared characteristics of these studies included 1) use of mild shock as a US, 2) an acquisition phase in which a CS+ and CS- are

⁵ It should also be noted that at this time, two analyses of a subset of the sample collected for this dissertation have been published (Hunt, Cooper, Hartnell, & Lissek, 2017, 2019). These manuscripts provide valuable insight into the discussed topics, but cannot be said to represent separate empirical endeavors or independent samples from the current study and are therefore excluded from this literature review.

conditioned and used as anchors on a generalization continuum (GSs) introduced in a subsequent phase, 3) avoidance was operationalized as active avoidance in the form of a behavioral decision (single button press) that completely negated the chance of shock in at least one portion of the experiment (the exception is that one of two related experiments in Arnaudova et al., 2017, Experiment 1, only includes a task that uses a response time index of avoidance that differs slightly from the other studies), and 4) measurements included both physiological and behavioral indices of fear and risk perception. All of the cited studies found evidence of generalized fear and generalized avoidance separately (i.e., not APIC-G), with generalization gradients indicating participants generally feared and avoided the closest approximation to the CS+ (i.e., the GS most similar to the CS+) more than the CS- or a control safety stimulus (e.g., a triangle in a task that uses circles for a GS continuum). Additionally, three of the four cited studies tested APIC-G as part of their analytic plan (the study by Arnaudova et al., 2017, did not include a correlational analysis or similar that directly tests APIC-G). In these studies, significant correlations between behaviorally indexed Pavlovian generalization (e.g., online risk ratings) and generalized instrumental avoidance ranged from $r = .46$ to $r = .83$, depending on how generalization was operationalized and if Pearson (generalization stimuli coded as continuous) or Spearman (generalization stimuli coded as ordinal) correlations were calculated. Only van Meurs et al. (2014) found a significant relationship between a physiological index of Pavlovian generalization (fear-potentiated startle measured by electromyography [EMG]) and generalized avoidance, with correlation coefficients ranging from .32 (association between overall shapes of Pavlovian and instrumental generalization) to .39 (generalization as indexed only by

response values for the closest approximation to the CS+ and no other stimuli). Although Boyle et al., (2016) and Cameron, et al., (2015) did not find significant associations between physiological indices of Pavlovian and instrumental generalization, both studies found non-significant correlations in the same direction (i.e., positive) and of very roughly comparable magnitude (r s ranged from .16 to .24). Differences between these studies are possibly due to 1) van Meurs and colleagues (2014) use of fear-potentiated startle to physiologically assess Pavlovian generalization compared with the use of skin conductance response (SCR) in the other two studies, which has been shown to be more susceptible to non-emotional influences and other sources of noise, and is perhaps less suited to generalization assessment than startle (Cacioppo, Tassinary, & Berntson, 2007; Lang, Bradley, & Cuthbert, 1998; Lissek, Biggs, et al., 2008), and 2) a larger sample size in van Meurs et al. (2016) compared with one of the other studies (44 in van Meurs et al., 2016, compared with 28 in Boyle et al., 2016). Differences in study design preclude more complex inferences from this small sample of studies, but the general conclusion that Pavlovian generalization is associated with instrumental generalization is supported, and evidence suggests that fear of a stimulus is contributing to avoidance of that stimulus and that these paradigms are suitable tools for probing this APIC-G.

Notably, only one of the cited studies (van Meurs et al., 2016) involves an explicit focus on the “approach” part of the approach-avoidance conflict: in this study, the participant completes a video game-like task that is framed in the context of a farmer harvesting crops; participants are explicitly instructed that approach results in a successful harvest, avoiding results in a high likelihood of losing the crops. The other cited studies do not provide explicit approach motivation, which might hinder

interpretation of results in the context of adaptive vs maladaptive avoidance decisions. To help illustrate the importance of including approach motivation as a variable in conceptualizes of APIC-G and relative adaptiveness of an APIC-G outcome, we return to the example of the child and the Rottweiler. As previously established, it is adaptive for the child to avoid the dangerous Rottweiler (CS+). That avoidance will be negatively reinforced as the child continues to successfully negate danger through avoidance, and this will keep the child safe regarding this particular Rottweiler. However, as the child generalizes fear and avoids safe dogs that only superficially resemble the dangerous Rottweiler (GS) they will also likely avoid those dogs, which are benign and could perhaps confer benefits to the child (e.g., companionship, increased confidence in ability to handle fear) and avoidance of them could result in interference in valued areas of living (e.g., being late to school, not wanting to go outside to meet friends) above and beyond what might be expected if the child was just avoiding the Rottweiler. In other words, the generalization of Pavlovian fear has led to generalization of avoidant behavior with negative consequences to the individual's functioning. Further, continued avoidance of safe dogs might result in lost opportunities for the child to extinguish their fear and unnecessarily negatively reinforces the avoidant behavior, which then continues to interfere with the child's goals and reward obtainment. Overall, it is clear from this example that the child's avoidance has become maladaptive, as avoidance has led to a decrease in behavioral repertoire and subsequent loss of reward that does not correspond to an actual increase in safety or negation of danger. In other words, the approach-avoidance conflict has been resolved with a non-optimal outcome.

A final consideration is how APIC-CS+ might fit into the approach-avoidance/adaptive vs. maladaptive framework. As the CS+ represents a genuine threat, it would be generally adaptive to avoid the CS+ and therefore APIC-CS+ can be considered an adaptive learning process. This was previously framed in the ongoing example as avoiding a dangerous Rottweiler is an adaptive choice. We can also view the inverse of APIC-CS+, approach behavior in the presence of the CS+, as maladaptive: a different child who continues to approach the dangerous Rottweiler and receives bites on every approach is not benefitting from additional learning opportunities (i.e., learning to avoid through instrumental learning) and is harmed every time they approach. That said, it should also be noted that uniformly deeming APIC-CS+ as adaptive might be oversimplifying – whereas the APIC-G clearly is related to unnecessary loss of reward or exhaustion of resources in the absence of actual threat or risk, APIC-CS+ might sometimes be maladaptive depending on context. For example, if we change our example slightly and assume that the child's parent has had the same learning experiences as the child and is as equally afraid of the Rottweiler, and is confronted with a situation in which the child is about to be bitten by the dog, a more complex approach-avoidance conflict is present. The parent presumably prioritizes their child's safety over the majority of all other concerns, and non-avoidance (intervening to save the child) in this case is considered adaptive because of this priority. However, for the purposes of this endeavor we will assume APIC-CS+ is generally adaptive, given that it takes a somewhat contrived or unusual situation to produce an example of adaptive CS+ approach and that in the vast majority of conditioning studies there is no analogue to the previously described example (for review of conditioning paradigms, see Lonsdorf et al., 2017).

In terms of empirical support for APIC-CS+ in humans, this exists both in the cited APIC-G studies and, more substantially, in differential fear conditioning studies that do not contain a generalization component (e.g., Alarcón, Bonardi, & Delamater, 2017; Claes, Vlaeyen, & Crombez, 2016; Delgado, 2009; Garofalo & Robbins, 2017; Lewis, Niznikiewicz, Delamater, & Delgado, 2013; Y. Xia, Gurkina, & Bach, 2019). Findings are consistent across these types of studies, with the CS+ related to an increased rate of instrumental avoidance (e.g., Delgado, 2009; Garofalo & Robbins, 2017; Lewis, Niznikiewicz, Delamater, & Delgado, 2013). The exception to this occurs in the minority of studies that contain an approach-avoidance conflict by including some form of reward manipulation (e.g., Alarcón, Bonardi, & Delamater, 2017; Claes, Vlaeyen, & Crombez, 2016), such as a graphical representation of reward (e.g., food) or winning points. These studies sometimes find that reward-seeking behavior “overrides” the instrumental avoidance learning, as some participants perhaps prefer to win regardless of a harmful outcome (e.g., Claes, Vlaeyen, & Crombez, 2016). These studies also typically have smaller sample sizes, which potentially creates vulnerability to unstable effects introduced through sample idiosyncrasies that might not be observed in another sample. Also of note related to APIC-CS+ is its clinical relevance to alcohol and drug dependency (Garbusow et al., 2016; Holroyd, Hajcak, & Larsen, 2006; Schad et al., 2019), as those with substance-related pathology will approach their desired substance even in the context of severe consequences (corresponding to maladaptive approach, as well as maladaptive APIC-CS+). Although this is not directly related to the current dissertation, it is important to note for later general discussion of reward-related personality traits.

Taken together, we see that both APIC-G and APIC-CS+ are clinically relevant, that significant effects for both have been documented in the literature, and that experimental procedures are sufficiently able to elicit these processes in the lab. Moving beyond the laboratory view, we return to our “real-life” example of the child and the Rottweiler. It is at this point that we might consider a question that has lingered in the background of this ongoing example: how does this child’s behavior compare to others, and is this particular child’s pattern of behavior normative or is it increasingly idiosyncratic and potentially problematic? Further, does it perhaps meet criteria for a psychological disorder? If taking a broader view, are there dispositional traits that contributed to or buffered against the development of the potentially pathological or maladaptive behavior? What information about the child’s dispositional traits and its relation to his behavior would be helpful to know before declaring an outcome adaptive or maladaptive? Overall, we are left with the essential question of “why did this child develop this set of behaviors, and others did not?”. The following section addresses these questions and related areas of inquiry.

Normative and Pathological Individual Differences in Generalized Fear, Generalized Avoidance, and APIC-G/APIC-CS+

Remarkably, only one previous study with a primary focus on APIC-G has tested for sources of potential inter-individual variance that might help sharpen prediction and improve explanatory accounts of the process (van Meurs et al., 2014), and even in this study only the instrumental avoidance outcome was correlated with a limited set of negative affect related traits and states. Further, this and the other previously reviewed studies do not find Pavlovian generalization to be highly predictive of instrumental

avoidance (i.e., the APIC-G relationship is of moderate strength), which is at odds with the intuition that people avoid feared stimuli because they are afraid of said stimuli and somewhat surprising given that this relationship is defined and constrained by experimental conditions that are nominally within the investigators control. That said, there are multiple reasons to view this as an unsurprising result. The explanation that is the focus of this dissertation is that what is actually viewed as error or “noise” in APIC-G and obscuring the true signal is actually inter-individual variation that can be measured and analyzed to help resolve the noise and produce a cleaner signal (see Lonsdorf & Merz, 2017, for a compelling and comprehensive review of this issue that we will discuss at multiple points throughout this dissertation).

The importance of inter-individual personality differences.

Inter-individual differences (hereinafter shortened to individual differences; intra-individual differences will be referred to as such or by specific description when needed) refers to the natural variation of people in a variety of domains, and further classification, delineation, prediction, and intervention related to individual differences and their associated behaviors is seen as a core tenet of psychological science (Eysenck, 1950; Eysenck & Eysenck, 1987; Harkness & Lilienfeld, 1997; Meehl, 1972, 1995; Paunonen & Ashton, 2001). Individual differences in human decision-making processes that underlie any behavior, including avoidance, can emerge from a vast range of possible sources of inter-individual variation (Aczel, Bago, Szollosi, Foldes, & Lukacs, 2015; DeYoung, 2015). Further, the decision-making process that underlies avoidance does not just receive input from neural fear circuitry (e.g., Spielberg, Miller, Warren, Engels, Crocker, Banich, et al., 2012; Spielberg, Miller, Warren, Engels, Crocker, Sutton, et al.,

2012; Tovote et al., 2016). Put simply, there are other sources of variation that contribute to why or why not an individual would avoid a feared stimulus besides variation in the fear variable itself, likely even if we concede that some of these sources of variation exert influence on avoidance through modulation of the initial fear itself. Past investigations of APIC-G, as well as a good portion of the fear conditioning literature, have unintentionally treated those sources of variation as noise. We now turn our attention to personality variation as one form of individual difference that can help us refine this noise and convert it into signal.

At its core, personality science and its constructs and techniques help us explain divergent outcomes between people in similar contexts, both those that are challenging and those that are routine (e.g., L. A. Clark, Watson, & Mineka, 1994; DeLongis & Holtzman, 2005; Watson & Clark, 1992; Watson, Clark, & Harkness, 1994; Webster & Ward, 2011). Personality taxonomies, in a sense, provide a framework through which to view these divergences and consider possible explanations, some quite different or unrelated to each other (e.g., John & Srivastava, 1999; Morey, Gunderson, Quigley, & Lyons, 2000; Shackman & Fox, 2018; Tackett, Quilty, Sellbom, Rector, & Bagby, 2008). This can be illustrated using a modification to the ongoing example: consider if there are now two children that have equally overgeneralized fear of dogs due to a harmful encounter with the Rottweiler. One of these children might avoid all dogs that resemble the dangerous Rottweiler, while the other child might choose not to avoid at all. Multiple factors could explain this difference: perhaps the non-avoidant child is more motivated by the potential positive value of interacting with dogs despite negative experience, or they cannot go to a birthday party without encountering the dog and value the party enough to

not avoid. Another explanation could be that the non-avoidant child has received firm instructions from their parents and would prefer to confront a dog than disappoint their parents. A third possible explanation is that the child is embarrassed by his fear and has resolved to aggressively confront the dog, even if it results in further harm. Hypothetical explanations are near limitless, but what is clear from this example is that if one asks the question “why do some people avoid and some do not?”, an answer that only refers to fear or the determinants of fear and no other possible factors is likely an incomplete answer and ignores individual variation that might contribute to APIC-G. Put another way: Pavlovian conditioned fear is, of course, relevant to the question of why people avoid, but it is not a complete answer. Interestingly, this and related questions are ones that the applied clinical literature has long considered relevant, as psychological treatment for anxiety pathology based on combinations of specific individual differences (i.e., personalized or individualized treatment; e.g., Ball, Stein, & Paulus, 2014; Cuijpers, Ebert, Acarturk, Andersson, & Cristea, 2016; Moscovitch, 2009; Ozomaro, Wahlestedt, & Nemeroff, 2013) or variation on a candidate personality dimensions (e.g., neuroticism or negative affectivity; Barlow et al., 2010; Barlow, Sauer-Zavala, Carl, Bullis, & Ellard, 2013; Sauer-Zavala et al., 2012) continues to be a prominent goal in the clinical science field. However, the question of personality individual differences in APIC-G has little to no footprint in the mechanism-focused experimental conditioning literature.

Review of individual differences work in generalization and related studies.

To outline the impetus for the current study and to help structure our hypotheses, we temporarily take a step back from a personality focus and in the following sections we review multiple types of individual differences in APIC-G, and, given the paucity of

relevant work in APIC-G (and APIC in general), experimentally-elicited conditioned fear generalization and instrumental avoidance generalization. The following section is organized into two broad classifications: normative individual differences (those that are mostly or entirely normally distributed in the population) and pathological individual differences (those that display a markedly skewed distribution in the population, with the majority of people tightly clustered around the lowest/absent values). This distinction is made partially to facilitate a natural organizational structure in both this section and the reported analyses, and to help consider which findings might be replicated in or relevant to the current study, which does not sample explicitly from a pathological or patient population (e.g., psychiatry clinic) and is therefore more likely to be representative of normative functioning and dispositional variation. We would also like to emphasize that this is an imperfect and partially artificial distinction, as copious research demonstrates both that normative individual difference scales (e.g., Big Five personality scales) are appropriate for measurement of clinical/pathological levels of trait variation (Kotov, Gamez, Schmidt, & Watson, 2010; Mahaffey, Watson, Clark, & Kotov, 2016; Malouff, Thorsteinsson, & Schutte, 2005; L. C. Morey et al., 2002, 2000; Rector, Bagby, Huta, & Ayearst, 2012; D. Watson & Naragon-Gainey, 2014) and that joint factor-analytic and other latent modeling approaches to the study of measures targeting normative and pathological personality variance find evidence that these spectrums actually exist on a single unified dimension (DeYoung, Quilty, & Peterson, 2007; Forbes et al., 2017; Kotov et al., 2017; Krueger & Tackett, 2003; Markon, Krueger, & Watson, 2005). Nevertheless, we contend it is a helpful framework for the current endeavor and will consider the ramifications of this distinction on our results and interpretation in the discussion.

Pathological individual differences.

Abnormalities in Pavlovian fear conditioning are a central part of etiological accounts of categorically-defined anxiety disorders (Dunsmoor & Paz, 2015; Lissek & van Meurs, 2014; Mineka & Zinbarg, 2006). One of the most fundamental abnormalities seen in anxiety disorders is overresponding to safety cues (Duits et al., 2015; Lissek et al., 2005), which provided the impetus for more recent studies of fear generalization in anxiety disorder patient populations using laboratory paradigms. The clinical relevance of fear generalization and, potentially, APIC-G is also established in the clinical literature. This is most directly evidenced by the fact that the DSM, which has been the guiding system for psychopathology research for a considerable amount of time, requires both fear or anxious distress and some form of avoidance be present for any of the most common anxiety, trauma, or obsessive-compulsive disorders to be diagnosed (American Psychiatric Association, 2013). The most common disorders of this type include posttraumatic stress disorder (PTSD), panic disorder, generalized anxiety disorder (GAD), specific phobia, social anxiety disorder (SAD), and obsessive-compulsive disorder (OCD) (Kessler, Ruscio, Shear, & Wittchen, 2009). The only potential exception to this consistent pattern across these DSM diagnostic criteria is GAD, although it can be argued that the required worry component of the disorder represents a cognitive form of avoidance that is central to GAD pathology and functionally resembles the more overt avoidance included in the DSM entries for the other listed disorders (Behar, DiMarco, Hekler, Mohlman, & Staples, 2009; Borkovec, Alcaine, & Behar, 2004). To facilitate

readability, specific findings for each of these disorders are discussed separately⁶. Given the field's prominent history of conditioning models for explaining the etiology of anxiety pathology, each disorder is briefly described and contextualized within the conditioning literature, followed by a review of relevant studies. When possible, pathological individual differences that were tested along a continuum, as opposed to categorically, are highlighted.

PTSD.

Formerly considered an anxiety disorder as recently as DSM-IV-TR (American Psychiatric Association, 2000) and now under the “trauma and stress-related disorders” heading of DSM-5, PTSD (American Psychiatric Association, 2013), PTSD is characterized by chronic fear and avoidance related to reminders of one or more past traumas accompanied by hyperarousal and negative mood and cognition symptoms. PTSD is the only DSM anxiety or trauma disorder with a clear precipitant (a “Criterion A” trauma, which most commonly for people diagnosed with PTSD is a direct exposure to actual or threatened death, bodily harm, sexual violence; e.g., Breslau, 2009) required for diagnosis, and therefore conditioning models are particularly intuitive and well-suited to characterizing this disorder due to clear links between the conditioning constructs and PTSD diagnostic features, both in terms of acquisition of fear (e.g., the threat during the trauma is a US, the cues present during the trauma comprise the CS+) and difficulty

⁶ Despite the inclusion of PTSD, acute stress disorder (ASD) is excluded from the current review. This diagnosis created to capture maladaptive functioning and intense distress in relation to trauma before a PTSD diagnosis can be made (i.e., < 1 month since trauma), there is considerable debate as to the validity of ASD as a diagnostic entity (Bryant et al., 2017; Marshall, Spitzer, & Liebowitz, 1999; McNally, 2003). Given the lack of fear generalization and instrumental avoidance studies with ASD, as well as the uncertainty regarding the diagnosis, we considered it a reasonable exclusion.

reducing the fear response over time (e.g., slower or delayed extinction learning). Proposed conditioning accounts of PTSD offer a compelling overview of how conditioning mechanisms contribute to the development and maintenance of PTSD (Bleichert, Michael, Vriends, Margraf, & Wilhelm, 2007; Lissek & van Meurs, 2014; Mahan & Ressler, 2012; Pitman, 1988; VanElzakker, Kathryn Dahlgren, Caroline Davis, Dubois, & Shin, 2014), particularly when viewed through the lens of two-factor theory. It should be noted that although most of these accounts are comprehensive in their explanations of the pathogenesis of PTSD and how it results in its constituent symptoms, both the authors of these accounts and other evidence suggests that there are factors other than conditioning mechanisms that contribute to the pathology, particularly in terms of risk factors for PTSD, such as reward- and impulse-related differences (e.g., Admon, Milad, & Hendler, 2013; Kramer et al., 2016). That said, these risk factors can be viewed as interacting with the conditioning aspects of PTSD (e.g., some risk factors contribute to enhanced fear conditionability, which in turn leads to greater chance of developing PTSD), and overall that conditioning models of PTSD perhaps offer the most explanatory power of current theoretical models, even with the noted gaps.

Fear generalization has been identified as a particularly important conditioning correlate of PTSD given the phenomenology of the disorder (Lissek & Grillon, 2012; Lissek & van Meurs, 2014; Zuj & Norrholm, 2019), and clinically-informed accounts of PTSD highlight the transfer of fear from the trauma itself to resembling cues as a core symptom and a priority treatment target (Ehlers & Clark, 2000; Foa, Steketee, & Rothbaum, 1989). At present, three studies have directly tested fear generalization in participants diagnosed with PTSD (Kaczurkin et al., 2016; Lissek & Grillon, 2012; R.

A. Morey et al., 2015). Two of these studies (Kaczurkin et al., 2016; R. A. Morey et al., 2015) used neuroimaging to identify neural substrates associated with overgeneralization in PTSD, with both studies identifying correlates of overgeneralized fear (e.g., insula, implicated in interoception and prediction of aversive events, Paulus & Stein, 2006) and deficient generalization of safety learning (e.g., ventromedial prefrontal cortex; implicated in inhibition of fear; Phelps, Delgado, Nearing, & LeDoux, 2004). Results notably differed between these studies, as abnormalities in some candidate brain areas implicated in fear generalization, including the amygdala, thalamus, and hippocampus were associated with PTSD in one study, but not the other. Further, Kaczurkin and colleagues (2016) found behavioral evidence of overgeneralization in PTSD in the form of ratings of shock expectancy (i.e., risk ratings), whereas Morey and colleagues (2015) did not. These differences are potentially explained by 1) differences in the experimental paradigms used (e.g., use of emotional faces in one paradigm and neutral shapes in another) and 2) heterogeneity in PTSD (DiMauro, Carter, Folk, & Kashdan, 2014) that might not be obvious in these samples due to both the assessment technique used to assess PTSD (e.g. and relatively small sample sizes to detect reliable signals of heterogeneity within the PTSD groups. Most importantly for the current endeavor, Kaczurkin et al., (2016) conducted additional analyses in which PTSD symptom severity was quantified dimensionally using total scores derived from a well-validated interview of PTSD symptoms (Blake et al., 1995) and correlated with a single-score measure of generalization gradient steepness (with higher scores indicating more shallow slopes and therefore greater generalization). These symptom severity scores were significantly associated with greater generalization in the right anterior insula and the left

ventral hippocampus/amygdala, indicating that generalization positively covaries with symptom severity level and not just with categorically-defined PTSD. Also noteworthy is that this study had a representative range and a normal distribution of symptom severity scores, as it tested participants categorized into trauma-control, sub-threshold PTSD, and PTSD groups; this suggests that these results reflect a true dimensional relationship in the population.

In terms of instrumental avoidance, generalized or otherwise, there are remarkably no published studies investigating this process in PTSD. This is particularly striking considering avoidance is one of the DSM-5 PTSD symptom clusters and required for diagnosis, that it is central to many etiological conceptualizations of PTSD (e.g., Asmundson, Stapleton, & Taylor, 2004; Solomon & Wynne, 1954; Thompson & Waltz, 2010), and that it is a primary intervention target for psychotherapeutic treatment (S. A. Rauch, Eftekhari, & Ruzek, 2012; Rothbaum et al., 2000). Even if expanding our review to experimental studies of avoidance in PTSD, there is still a limited amount of extant relevant research to help inform the current study. One study used a video-game task that allowed participants to either engage for points or hide to escape aversive on-screen outcomes and point loss (Sheynin et al., 2017). These studies identified a consistent pattern of more persistent avoidance behavior in a high PTSD symptom group, as indexed by avoidance during a phase in which it was not possible to lose points or experience the aversive event. Two limitations hinder interpretability of the results. First, the findings purportedly support a link between lab-based avoidance and avoidance symptoms of PTSD, but a plausible alternative interpretation is that those with higher PTSD symptoms were better at learning the task more likely to adaptively avoid and

maximize success on the task; it is also notable that the higher PTSD symptom group had significantly more points at the end of the task than the lower symptom group. Second, the instrument used to define the groups is an imperfect measure of PTSD and can lead to false positive classification of PTSD (Wilkins, Lang, & Norman, 2011); the authors address this in their study, but it is important to consider the possibility that these results are not particularly informative in regards to pathological individual differences. Finally, given the paucity of research around instrumental avoidance in PTSD, it is no surprise that there have not been any investigations of APIC-G and this remains a needed area of investigation.

GAD.

Intense and pervasive worry that is present for at least six months in conjunction with related anxious-arousal symptoms are required for a GAD diagnosis (American Psychiatric Association, 2013). The etiology of GAD is relatively less well-understood when compared with many of the other anxiety disorders, with multiple overlapping theoretical accounts of GAD development and maintenance in existence but with no clear dominating view (Behar, DiMarco, Hekler, Mohlman, & Staples, 2009). Limited conditioning models of GAD have been proposed and provide some explanatory power for GAD pathology, but they either are not comprehensive and/or are still in preliminary stages (Cooper, Grillon, & Lissek, 2018; Mineka & Zinbarg, 2006) or are currently limited to animal research (Delgado, Olsson, & Phelps, 2006; Luyten, Vansteenwegen, van Kuyck, Gabriëls, & Nuttin, 2011). The most compelling evidence for the role of conditioning abnormalities in GAD pathology is from fear generalization work done in this population. Three studies provide consistent evidence that people with GAD

overgeneralize fear compared with control participants (Cha et al., 2014; Greenberg, Carlson, Cha, Hajcak, & Mujica-Parodi, 2013; Lissek et al., 2014), with all results pointing towards GAD-specific increased fear reactivity to the GS that was the closest approximation to the CS+, and the two studies using neuroimaging (Cha et al., 2014; Greenberg et al., 2013) identifying deficient inhibitory learning as indexed by a lack of vmPFC response to GSs. However, two studies provide contradictory evidence, both from the same research group (Tinoco-González et al., 2015; Torrents-Rodas et al., 2013). Both do not find differences in fear generalization between an anxiety group and a control group. In the first study, a continuous measure of trait anxiety is dichotomized in high/low group and proposed as an analogue of GAD; in the second study GAD is assessed using a brief diagnostic interview and with the same measure of trait anxiety. A possible explanation for the difference in results between these two studies and the three studies with significant generalization findings lies in these grouping details. In both studies with null results, the mean trait anxiety was considerably lower than both the mean trait anxiety level in the other studies and the recommended diagnostic cutoff for the measure (e.g., Kvaal, Ulstein, Nordhus, & Engedal, 2005). Thus, it is possible that the GAD group was miscategorized and that it and anxiety disorder analogue group were not sufficient for group differences to emerge.

At present, there are no published studies of instrumental avoidance, generalized instrumental avoidance, or APIC-G in GAD. This is less surprising than the lack of literature on this topic for PTSD, as there is still debate over the role of avoidance in GAD (Behar et al., 2009; Olatunji, Moretz, & Zlomke, 2010), especially overt behavioral avoidance, and there is not as compelling a case for study of instrumental avoidance in

GAD as there is for PTSD. However, normative dimensions of negative affect, such as Neuroticism and trait anxiety, are considered in the following section, and the review of those dimensions as they relate to instrumental avoidance has some relevance to GAD given conceptual and empirical overlap (see Kotov, Gamez, Schmidt, & Watson, 2010; Kotov et al., 2010; Watson, 2005).

Panic disorder.

Panic disorder is defined by a history of uncued panic attacks, in which the person rapidly experiences intense physiological symptoms and fearful cognitions, and subsequent intense worry or behavioral avoidance related to fear of having another panic attack (American Psychiatric Association, 2013). Conditioning accounts of pathological panic acquisition are prominent in the etiology literature (Acheson, Forsyth, Prenoveau, & Bouton, 2007; Bouton, Mineka, & Barlow, 2001; De Cort et al., 2017; De Cort, Griez, Büchler, & Schruers, 2012) and, broadly, posit that Pavlovian fear conditioning occurs during the initial panic attack, in which the panic symptoms are the US and incidental cues in the environment become CSs⁷. Further, in many cases, people with PD develop a pathological aversion to and avoidance of settings and cues which they associate with panic attacks and believe they would have difficulty escaping from (i.e., agoraphobia). Although not required for a PD disorder, agoraphobic symptoms are evident in ~25% to 50% of PD cases (Grant et al., 2006; Kessler et al., 2006) and are central to widely accepted models of the disorder and its treatment (e.g., Barlow, Craske, Cerny, & Klosko, 1989; Craske & Barlow, 1988).

⁷ Although beyond the scope of the current dissertation, it should be noted that there is some debate and contention around what constitutes the US, CS+/CS-, and other equivalents to conditioning phenomena in panic disorder (for discussion, see Bouton, Mineka, & Barlow, 2001; McNally, 1990).

As with PTSD, fear generalization has been proposed as a compelling explanation for how panic disorder pathology develops and maintains over time (Duits et al., 2016; Lissek et al., 2009). At present, one study demonstrates fear overgeneralization in panic disorder compared with control participants (Lissek et al., 2010); there are no equivalent studies that use a method appropriate for dissociating the separate roles of fear/excitatory and safety/inhibitory learning (e.g., fMRI).

Experimental avoidance studies of panic disorder are also limited, as the majority of studies of panic-related avoidance involve correlating responses to an acute stressor (e.g., CO₂ inhalation) with psychometric self-report indices of avoidance behaviors (e.g., Bystritsky, Craske, Maidenberg, Vapnik, & Shapiro, 2000; Spira, Zvolensky, Eifert, & Feldner, 2004). Thus, there are currently no relevant instrumental avoidance studies available for review. One study that potentially has implications for APIC-G used an idiographic anxiety induction and compared panic disorder patients who used in-situation safety-seeking behaviors (e.g., to those who did not, and found that those who did not use safety-seeking behaviors (which were sometimes overt avoidance and sometimes more covert forms of avoidance, such as emotional grounding using a held object; e.g., Funayama et al., 2013) reported significant decreases in negative beliefs and anxiety (Salkovskis, Clark, Hackmann, Wells, & Gelder, 1999). Although an imperfect analogue, this finding does roughly correspond to the assertion that generalized fear (in this study, fear related to learned threat cues) is associated with generalized avoidance (in this study, safety-seeking behaviors in response to these cues) and that, consistent with two-factor views, this serves to maintain the learned fear associations and facilitates continued anxiety.

SAD.

People with SAD report severe and pervasive anxiety in relation to social activities, either generally or in specific areas (e.g., public speaking), and pathologically avoid feared social contexts due to concern regarding the judgments and evaluations of others (American Psychiatric Association, 2013). Etiological accounts of SAD trend towards cognitive models, in which different forms of interpretation biases (e.g., tendency to interpret interpersonal information as negative) are posited as key factors in the development and maintenance of SAD symptoms (D. M. Clark & Wells, 1995; Hofmann, 2007; Rapee & Heimberg, 1997; Spence & Rapee, 2016). However, conditioning abnormalities are also considered part of SAD etiology: over-conditionability and impaired discrimination learning related to threatening social cues have been identified in patients with SAD (Hermann, Ziegler, Birbaumer, & Flor, 2002; Lissek, Levenson, et al., 2008; Sachs, Anderer, Doby, Saletu, & Dantendorfer, 2003), and conditioning mechanisms have been incorporated into cognitively-focused causal models that establish how a “landmark” stressful events or series of events (e.g., bullying) lead to threat cues that are part of the development of negative interpretation biases (Hofmann, 2008; Mineka & Zinbarg, 2006; Rapee & Spence, 2004; Spence & Rapee, 2016).

Currently, there is one published study that has found evidence for fear overgeneralization in SAD (Ahrens et al., 2016). Notably, this study did not use shock as a US, as was used with the majority of other generalization studies; instead socially-relevant USs were used (co-occurring scream and fearful face) which is in line with other studies that have used conditioning paradigms with social USs to differentiate those with SAD from healthy control participants (e.g., Lissek, Levenson, et al., 2008) and suggests

US specificity plays a role in fear generalization (i.e., for generalization to occur the US must be disorder relevant). Also notable is that Ahrens and colleagues (2016) found physiological, but not behavioral, generalized conditioned responses to the CS+ and the three classes of GSs must similar to the CS+ in SAD, but not controls. Behaviorally, they found overall greater threat estimation in SAD when compared with control participants (i.e., a main effect), but no evidence of generalization differences. This is perhaps reflective of the prominence of cognitive, as opposed to emotional or physiological, factors underlying excessive and pervasive apprehension of social stimuli in those with SAD (Spence & Rapee, 2016); although it should be noted that desynchrony between physiological and behavioral measures of generalization has been documented in other disorders (Greenberg et al., 2013) and is consistent with earlier theories of these disorders (Hodgson & Rachman, 1974).

Approach-avoidance conflict in SAD has received moderate empirical attention, despite SAD and its signature conflict between social fear and valued activities being an intuitive example for understanding how approach-avoidance conflicts operate outside of the laboratory setting (Pittig, Treanor, et al., 2018). Although there are no APIC-G in SAD studies published as of this date, two studies have experimentally investigated covariation of fear (experimentally conditioned and unconditioned) and instrumental avoidance in SAD (Ly & Roelofs, 2009; Pittig, Alpers, Niles, & Craske, 2015⁸) and found somewhat contradictory instrumental avoidance findings. Ly and Roelofs (2009)

⁸ Although not explicitly conceptualized as an instrumental *conditioning* task, the paradigm used in Pittig et al. (2015) does meet the definition for instrumental avoidance, as the behavioral outcome (gambling choice) was shaped through prior experience. The first author also chose to categorize this study as one of instrumental avoidance in a later review he wrote (Pittig et al., 2018).

compared those with high SAD symptoms to those with low SAD symptoms and did not find differences in instrumental avoidance rates between the two. Pittig and colleagues (2015) found a positive correlation between preconditioned (i.e., *in vivo* or naturally conditioned) fear and avoidance decisions in a group of SAD patients completing a public speech and a gambling task. Numerous differences in study design and population could account for these differences, most pertinently that 1) one study only assessed aversive avoidance without a reward component (Ly & Roelofs; 2009) while the other study (Pittig et al., 2015) included both fear and reward but they were part of separate tasks that were only statistically, not experientially, related and 2) one study used SAD analogues that perhaps are not representative of SAD patients and the other only conducted within-group analyses of SAD patients and did not include a control group. Given these complications, it is difficult to make strong conclusions about APIC-G in SAD based on these data, and further investigations are needed.

Specific phobia.

Specific phobia involves excessive fear and avoidant behavior in response to a narrowly-defined stimulus or context, such as spiders, blood and needles, or confined spaces (American Psychiatric Association, 2013). Due to the circumscribed nature of fear in this disorder and relative homogeneity of symptoms within specific phobia types, conditioning models are seen as the most parsimonious explanation for the pathology and have held dominance in the field for decades (Field, 2006; McNally, 1987; Mineka & Zinbarg, 2006; Stein & Matsunaga, 2006), with a sizable number of studies documenting fear conditioning abnormalities in different types of specific phobias (see Duits et al., 2015) and using this evidence base to develop remarkably effective and brief treatments

for the condition (Choy, Fyer, & Lipsitz, 2007; Öst, 1989, 1996). However, it should be noted that although conditioning models are the dominant view of social phobia etiology, this does not mean that fear acquisition/simple discrimination learning models are dominant; current theories posit that multiple conditioning mechanisms contribute to the specific phobia (e.g., vicarious conditioning; Field, 2006; Mineka & Zinbarg, 2006). This is important to note because it might seem that there is no need to go beyond simple fear acquisition models if considering specific phobia from the perspective that it is based on a simple link between a US (e.g., dog bite) and CS+ (e.g., the dog).

Indeed, fear generalization is also implicated as a mechanism underlying specific phobia pathology. To date, two studies have experimentally investigated conditioned fear generalization in specific phobia (Dymond, Schlund, Roche, & Whelan, 2014; Lange et al., 2019). Both studies compared participants with spider phobia to non-phobic controls. Only one study, Dymond et al. (2014) found those with spider phobia demonstrated symbolic (as opposed to perceptual) overgeneralization (in the form of behavioral ratings of threat expectancy). The other study conducted by Lange and colleagues (2019) found support for enhanced conditionability in specific phobia (stronger responses to the CS+) and enhanced (not impaired, as found in prior studies) discrimination learning. The difference in results might be due to Dymond et al. (2014) using a symbolic generalization task (i.e., part of the task involved learning arbitrary relationships between dissimilar stimuli), as opposed to the perceptual generalization tasks used in most other studies, including Lange et al. (2019), and suggests that perhaps those with specific phobia are more susceptible to generalization of newly learned arbitrary associations, but can successfully discriminate between the visual properties of the CS+ and a close

approximation. Additionally, Dymond et al. (2014) used a disorder-specific US (spider images), whereas Lange et al. (2019) used a disorder-agnostic US (shock), which potentially created a strong situation that obscured group differences (i.e., elicited a uniform defensive response across all participants, see Lissek, Pine, & Grillon, 2006).

Instrumental avoidance has been directly studied in two investigations of specific phobia (Pittig, Brand, Pawlikowski, & Alpers, 2014; Rinck et al., 2016), both of which found greater avoidance in participants with spider phobia than controls. However, interpretation and implications differ based on task design: Pittig et al. (2014), which contained an approach-avoidance manipulation (modified gambling task) frame the finding of increased avoidance as maladaptive because it came at the cost of potential reward, whereas Rinck et al. (2016), which only contains aversive stimuli with no reward or approach element, frame their results as more efficient learning of when to avoid a spider. Although initially appearing to contradict each other, these findings are actually quite compatible: because Rinck et al. (2016) did not contain a reward/approach element, there is not proper context to deem avoidance as adaptive or maladaptive. Regardless, it is clear that those with specific phobia are more likely to avoid feared stimuli. Neither study provides direct insight into fear and avoidance covariance or APIC-G, although both found that trait levels of spider fear were associated with more avoidance decisions. This is also consistent with a number of studies using a Behavioral Approach Task (BAT) that find greater avoidance in those with phobias, as indexed by fewer approach “steps” toward a CS+ compared with controls (Lange et al., 2019; Olatunji, Cisler, Meunier, Connolly, & Lohr, 2008; Valentiner, Telch, Petrucci, & Bolte, 1996; Zoellner, Echiverri, & Craske, 2000). These studies also perhaps provide insights into how experimentally-

elicited generalized avoidance might function in specific phobia, as the BAT perhaps has inherent generalization properties due to its design as a graded approach task that (e.g., first participants see the CS+ from a distance, then walk closer, and eventually touch the CS+), although it should be noted that sensitization (i.e., non-associative increase in a response to a stimulus over repeated exposures) might be a better explanation for this avoidance pattern than an associative process such as generalization.

OCD.

The cardinal features of OCD are distressing and repeated obsessive thoughts and compulsive behaviors enacted to control or reduce distress, typically distress related to the obsession. (American Psychiatric Association, 2013) As with PTSD, OCD was reclassified in DSM-5 and was moved from the anxiety disorder to the obsessive- and tic-related disorders section. It is included due to its mechanistic similarities to the other discussed disorders. Although there is support for a conditioning model of OCD etiology (Armstrong & Olatunji, 2017; Geller et al., 2017; Milad et al., 2013; Mineka & Zinbarg, 2006; Nanbu et al., 2010; Tracy, Ghose, Stecher, McFall, & Steinmetz, 1999), the primary contribution of conditioning principles to the OCD literature has been in the area of intervention, in which exposure and response prevention techniques have become a first-line treatment for OCD and include components related to two-stage theory (Abramowitz, 1996; Foa et al., 2007; Foa & McLean, 2016; Ludvik, Boschen, & Neumann, 2015). Further, the role of conditioning processes in OCD is complicated by the fact that the majority of conditioning studies involve conditioned fear, yet there is evidence that disgust, as opposed to fear, is the driving motivational force in many cases of OCD (Cisler, Olatunji, & Lohr, 2009) and that heterogeneity in the disorder might

reflect distinct phenotypes that might not be appropriately described with a single conditioning model (McKay et al., 2004).

As Pavlovian conditioning studies of OCD are limited (Duits et al., 2015; Lissek et al., 2005), it is not surprising there is only one published study of conditioned fear generalization related to OCD or obsessive-compulsive traits (Kaczkurkin & Lissek, 2013). In this study, limited evidence for psychophysiological generalization of conditioned fear was found for those with a higher, but not lower disposition towards threat estimation as measured by a self-report measure of obsessive beliefs. Accordingly, it is difficult to make substantial inferences regarding conditioned fear generalization in OCD based on this study.

Similarly, inferences regarding instrumental avoidance/generalized avoidance and APIC-G are difficult, as to our knowledge there are no available studies on these processes in diagnosed OCD. However, one study that tests an OCD analogue group in comparison to a control group using an instrumental avoidance test does provide some insight (Hassoulas, McHugh, & Reed, 2014). Results from this study, which contained both an un-signaled avoidance task (point loss was uniformly accomplished via a specific button press) and an instrumental avoidance task (cues indicated which combination of button-pressing responses would result in avoidance of point loss) indicated that those higher on OCD traits showed greater un-signaled avoidance (i.e., habitual avoidance) but did not differ in terms of instrumental avoidance. Also of note, the finding of increased un-signaled/habitual avoidance in OCD has been documented in previous studies (Gillan, Apergis-Schoute, et al., 2014; Gillan, Morein-Zamir, et al., 2014). Taken together, it is

possible that OCD is not related to instrumental avoidance abnormalities and is characterized by broad tendencies towards avoidance, but further research is needed.

Interim summary of pathological individual differences.

As there are inconsistencies in findings both across disorders and within disorders, a complete synthesis is not tenable. However, some specific conclusions are possible despite this heterogeneity of findings. First, there is strong support for overgeneralization of conditioned fear as a transdiagnostic feature of anxiety and related disorders, although the strength of this effect might vary by disorder. Second, a general tendency towards avoidance in experimental studies is seen across the disorders, although this is fairly unsurprising given the avoidance symptoms required for diagnosis, and verges on circular logic. Transdiagnostic conclusions for the role of experimentally-induced instrumental avoidance, generalized or otherwise, in the reviewed disorders are not possible due to 1) a lack of relevant published findings and 2) inconsistency in the findings that do exist (e.g., for specific phobia). However, theoretical conditioning models of these disorders based on clinical observation and treatment studies generally support the statement that APIC-G and its constituent processes are relevant to the etiology of anxiety/trauma/obsessive-compulsive pathology, with empirical support limited to inferences constructed from a set of findings that provide indirect evidence (for review, see Pittig, Treanor, et al., 2018). It is noteworthy that there are no available studies of dimensionally-conceptualized or measured pathological individual difference (e.g., clinical scales from the Minnesota Multiphasic Personality Inventory [MMPI] or Personality Assessment Inventory [PAI], see Butcher & Rouse, 1996) as they pertain to fear and avoidance generalization, despite clear relevance of these multiband scales to

fear conditioning processes (e.g., the subscales of the “ANX” scale on the PAI that capture cognitive vs. physiological symptoms of anxiety independent of specific disordered manifestations; L.C Morey & Boggs, 1991). Related to this, we also acknowledge the inherent heterogeneity in and comorbidity issues related to DSM disorders (Krueger, 1999; Krueger & Markon, 2006; Lilienfeld & Treadway, 2016) and that alternative empirically-derived classification structures (Kotov et al., 2017; Krueger, McGue, & Iacono, 2001) might better map onto the basic fear generalization substrates identified in the studies using DSM disorder.

Narrowband normative individual differences.

Normative/non-pathological individual differences in generalized fear, avoidance, and APIC-G have received less empirical focus than the study of pathological differences in these processes. This is likely due to multiple factors, including: 1) fear and avoidance conditioning are commonly linked to clinical conditions and hold intuitive appeal in how they relate to those conditions; 2) these individual differences are typically modeled as continua, which is difficult to incorporate into the ANOVA/factorial statistical framework that predominates the fear conditioning field (e.g., Lonsdorf & Merz, 2017; Vanbrabant et al., 2015); and 3) until recently, reliance on categorical DSM disorders as the predominant measurement model of experimental psychopathology (e.g., Cuthbert & Insel, 2013; Krueger & DeYoung, 2016)

Due to this lack of literature, we also briefly review normative individual differences in fear acquisition measured during discrimination learning paradigms when available, in addition to reviewing findings from generalization paradigms; this is justified because discrimination learning with a threat and safety cue or condition can be

conceptualized as a form of generalization (Lissek et al., 2005), just on the opposing pole of the same dimension (i.e., greater discrimination = less generalization). Additionally, the following review makes a further distinction between types of individual differences and is thus subdivided into “broadband” and “narrowband” individual differences or personality variables (Goldberg, 1999)⁹. In this case, broadband refers to superordinate traits that can be conceptualized as capturing coarser, but also wider, variance and content areas than narrowband traits (for examples of broader vs. more specific personality coverage, see Goldberg, 1990; Hofstee, de Raad, & Goldberg, 1992). Put another way, broadband traits have greater sensitivity and broader coverage of variance and content area, but less resolution and specificity. The opposite is true for narrowband traits, which are typically specific to a single content area or a closely related group of constructs.

Trait anxiety.

Trait anxiety is the individual tendency to experience high levels of anxiety in the moment (Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983). In other words, the higher a person’s trait anxiety, the more likely they are to respond anxiously or experience thoughts, feelings, and sensations associated with anxiety (i.e., with higher state anxiety). It is purportedly normatively distributed in the population and does not specifically demarcate pathological anxiety, although studies have identified cutoff scores for detection of pathology (e.g., Kvaal et al., 2005). The specific domain or type of

⁹ We conceptualize all the individual differences variables discussed in this review and tested in the current study as reflecting personality (and consistent with Tellegen (1991) and others’ definitions of personality) and therefore accurately termed personality variables; however, to facilitate the pathological vs. normative dichotomy, within this section we refrain from terming variables as personality variables unless they are part of an explicitly defined personality structure (e.g., Big Five).

anxiety is not considered in trait anxiety, which is one of the primary criticisms directed at the trait anxiety construct (Endler & Kocovski, 2001; Reiss, 1997). Another relevant criticism is that the trait anxiety construct is not sensitive to anxiety, but rather to negative affect in general, and does not reliably discriminate between anxiety and depression conditions (Bieling, Antony, & Swinson, 1998; Endler, Cox, Parker, & Michael, 1992; Kennedy, Schwab, Morris, & Beldia, 2001). Despite these limitations, it is the most common normative individual difference tested in association with fear and avoidance conditioning processes (Lonsdorf & Merz, 2017). In terms of fear acquisition, results are very mixed in terms of significant vs null findings, directionality or source of findings, and desynchrony between different measures of fear in the same study. Studies with significant results can be loosely categorized into those that found a relationship between trait anxiety and a poorer discriminative fear conditioning driven by reduced responsivity to the CS+ (Indovina, Robbins, Núñez-Elizalde, Dunn, & Bishop, 2011; Sjouwerman, Scharfenort, & Lonsdorf, 2018) and those that found this relationship was driven by heightened responsivity to the CS- (Gazendam, Kamphuis, & Kindt, 2013; Haaker et al., 2015; Haddad, Pritchett, Lissek, & Lau, 2012; Kindt & Soeter, 2014). Also, in most of the cited studies, there was some inconsistency between physiological and behavioral findings (i.e., one index would be significantly related to trait anxiety, but not the other). Additionally, as of this date, there are more null findings than significant findings in this sub-area (for a more in-depth review and discussion of this issue, see Lonsdorf & Merz, 2017).

Only a single study has explicitly focused on investigation of trait anxiety as it relates to conditioned fear generalization (Torrents-Rodas et al., 2013). As previously

discussed, the high trait anxiety group in this study was also proposed as a GAD analogue group, and no generalization differences were found, possibly due to limitations related to sampling or analytic technique. Similarly, the only published analyses of instrumental avoidance as it relates to trait anxiety did not find a significant relationship (Lommen, Engelhard, & van den Hout, 2010; van Meurs, Wiggert, Wicker, & Lissek, 2014). There are no published analyses of the relationship between trait anxiety and APIC-G (or APIC overall).

Due to substantial methodological and sample heterogeneity within these and the previously cited studies, most notably as it relates to 1) US reinforcement parameters (e.g., 100% vs < 100% reinforcement, which has been shown to substantially affect between- and within-subjects differences in fear responding, e.g., Chase, Kumar, Eickhoff, & Dombrowski, 2015; Chin, Nelson, Jackson, & Hajcak, 2016; W. Xia, Dymond, Lloyd, & Vervliet, 2017); 2) how the trait anxiety dimension was dichotomized to facilitate categorical between-subjects analyses (i.e., what cut-score was used, see Lonsdorf & Merz, 2017, p. 710, Figure 3, for a striking visualization of heterogeneity in chosen cut-score); and 3) sample size (a majority of the cited studies were susceptible to type II error due to being underpowered), we refrain from providing an overall interpretation of trait anxiety effects. A single exception is that findings of exaggerated fear responding to the CS-, which can be conceptualized as a form of GS, in those higher on trait anxiety (e.g., Gazendam, Kamphuis, & Kindt, 2013; Haaker et al., 2015; Haddad, Pritchett, Lissek, & Lau, 2012; Kindt & Soeter, 2014) is perhaps indicative of overgeneralization (or at least increased generalization) being associated with trait anxiety.

Intolerance of uncertainty.

Intolerance of uncertainty (IU) is a trait that captures a tendency towards negative beliefs and maladaptive behaviors regarding future events with uncertain outcomes and consequences (Buhr & Dugas, 2002; Dugas, Buhr, & Ladouceur, 2004). IU has more recently become of particular empirical interest given its can serve as an index of differential responding to uncertainty and ambiguity, which has been proposed as a behavioral correlate of activity neural circuits relevant for optimal decision-making and goal-selection (e.g., Grupe & Nitschke, 2011; Grupe, 2017; Sarinopoulos et al., 2010). It is moderately-to-strongly positively associated with trait anxiety (Khawaja & Yu, 2010; Sexton & Dugas, 2009) and has been identified as a core psychological substrate of multiple affective disorders (Carleton, 2012, 2016; Einstein, 2014; Shihata, McEvoy, Mullan, & Carleton, 2016), with a particular focus on how it relates to GAD and potentially facilitates the development of pathological worry that underlies the disorder (Deschênes, Dugas, & Gouin, 2016; Dugas et al., 2004; Dugas & Ladouceur, 2000).

IU is still a relatively new variable in terms of formal empirical investigation using conditioning techniques, and fear conditioning investigations of IU are limited. As with trait anxiety, findings are mixed and somewhat contradictory, but in this case, there are explanations that help to somewhat resolve these inconsistencies. First, in line with conceptualizations of IU as predicting sensitivity to uncertainty, significant effects for IU have been found when measuring fear acquisition to cues with uncertain signal value or measuring fear response in a period of experimentally-controlled uncertainty (Chin et al., 2016; Morriss, 2019; Morriss, Saldarini, Chapman, Pollard, & van Reekum, 2019; Nelson & Shankman, 2011). Unfortunately, these findings contradict each other in terms

of directionality and/or strength of the associations. To further complicate interpretation, other studies have not found any significant IU differences during unpredictable conditions or in response to unpredictable/ ambiguous cues (S. Chen, Yao, & Qian, 2018; Dunsmoor, Campese, Ceceli, LeDoux, & Phelps, 2015; Morriss, Saldarini, & Reekum, 2018; Morriss & van Reekum, 2019). Finally, another set of findings are consistent with predictions then when threat/danger is certain, IU effects will not be present (Morriss, Christakou, & Reekum, 2015; Morriss, Christakou, & van Reekum, 2016), although this also is somewhat contradicted by findings from Chen et al. (2018), who find that IU is associated with increased worry during a period of certain threat. Overall, these discordant results are possibly explained by the considerable methodological heterogeneity that is present in these studies and, as discussed in the previous section about trait anxiety, can have substantial and, problematically, idiosyncratic influences on results (Lonsdorf & Merz, 2017). In this case, likely culprits are differences in the US used (e.g., shock vs aversive tone), differences resulting from use of paradigms using cue-based conditioning vs those using context based-conditioning or negative affect eliciting procedures, variability in the IU self-report measure used, or differences in fear measurement method (e.g., SCR vs behavioral report). It should also be noted that although IU effects on fear acquisition are not consistent, there is somewhat more homogeneity in results during fear extinction, a process which potentially maps on the IU construct more cleanly (Morriss et al., 2018; Morriss & van Reekum, 2019).

At present, there is one study of IU and fear generalization, in which a negative relationship between Prospective IU (an IU subscale), and overgeneralization as indexed by electrocortical response (EEG measured event-related potentials [ERP]); a behavioral

index and IU were not related (Nelson, Weinberg, Pawluk, Gawlowska, & Proudfit, 2014). This finding is explained by noting that the ERP measured, the late positive potential, is sensitive to sustained attention towards affectively arousing stimuli, differs from common fear measures. Accordingly, this analysis potentially captured those with higher Prospective IU exhibiting a more adaptive discrimination process that results in attenuated generalization. There is also only one study of IU and instrumental avoidance (Flores, López, Vervliet, & Cobos, 2018). This study also found results with Prospective IU in particular, as it was associated with increased avoidance rate during the instrumental conditioning period. Taken together, the considerable heterogeneity of IU results leads to hesitation in making strong conclusions based on the existing data. At present, it appears that Prospective IU might be more sensitive to experimental conditioning manipulations and perhaps is related to fear generalization and avoidance, but further research is clearly needed. At this time, the main conceptual link between IU and heightened generalization is related to the assertion that the GSs communicate threat uncertainty. Also of relevance is that GAD has been linked to overgeneralized fear (e.g., Lissek et al., 2014), which might indicate IU would be related to fear generalization as well given the higher levels of IU observed in the disorder.

Anxiety sensitivity.

Anxiety sensitivity (AS) reflects the individual tendency to consider anxiety and fear related sensations as threatening and is sometimes termed “fear of fear” (Reiss, Peterson, Gursky, & McNally, 1986). As with IU, there is an established association between trait anxiety and AS (McNally, 1989; Sandin, Chorot, & McNally, 2001), as well as an association between AS and IU itself (Allan et al., 2017; Fergus & Bardeen,

2013). There has been some contention in the field around the construct validity of AS and if it represents a separate construct from trait anxiety (Lilienfeld, 1997; Lilienfeld, Jacob, & Turner, 1989; Lilienfeld, Turner, & Jacob, 1993) or if it is a highly-related but distinguishable construct (McNally, 1989; Peterson & Heilbronner, 1987; S. Taylor, 1996). The most consistent claim for AS as a separate construct is based around its specificity towards somatic and physiological sensations of anxiety or those related to anxiety (Joiner Jr et al., 1999; Ocañez, McHugh, & Otto, 2010; Schmidt & Joiner, 2002; Sturges, Goetsch, Ridley, & Whittal, 1998) and its differential relationship to anxiety disorders when compared with trait anxiety (Naragon-Gainey, 2010; Olatunji & Wolitzky-Taylor, 2009), with AS demonstrating particular relevance as a pathogenic mechanism contributing to panic disorder (McNally, 2002; Poletti et al., 2015; Schmidt, 1999).

At present, AS is the least studied of the reviewed narrowband normative individual differences in the context of fear conditioning, with only a single study relevant to the current review (Forsyth, Palav, & Duff, 1999). In this study, Forsyth and colleagues (1999) tested a trichotomized sample (high/medium/low AS) using a CO₂ inhalation for the US to specifically probe AS-related responses that were paired in a conditioning procedure with brief movie clips of negative or neutral valence. The authors did not find significant differences between AS groups on physiological indices, which is potentially explained by a confound in conditioning procedure: the CSs differed in valence and relevance to interoceptive vs exteroceptive anxiety sensations, which potentially activated a covariation bias (e.g., Alloy & Tabachnik, 1984; Tomarken, Mineka, & Cook, 1989) that was a stronger influence than the AS-specific manipulation

of CO₂ inhalation and thus obscured AS group effects. Another explanation lies in the recurring theme of artificial categorization of a continuous trait complicates statistical testing and inference, with this being a particular problem with trichotomization due to it introducing another potentially arbitrary cut point over median-split techniques (Zedeck, 1971).

To our knowledge, there are no available studies of AS and conditioned fear or instrumental avoidance generalization. In terms of other fear-or threat-related avoidance measured using other paradigms, a small selection of studies provide some insight, with significant findings of higher AS associated with greater avoidance of spider images (Lebowitz, Shic, Campbell, Basile, & Silverman, 2015) and risky-decisions (Broman-Fulks, Urbaniak, Bondy, & Toomey, 2014). There are also correlational and latent factor studies that find positive associations between AS and self-reported measures of behavioral avoidance (e.g., Asmundson & Taylor, 1996; Wilson & Hayward, 2006). Taken together, there is preliminary evidence that AS is related to overall avoidance of aversive stimuli and situations, but conclusions related to conditioned fear or avoidance, generalized or not, are not possible at this time. One possible exception is that it is reasonable to predict AS is associated with fear generalization based on findings of overgeneralized fear in panic disorder (Lissek et al., 2010).

Broadband normative individual differences.

Neuroticism.

One of the most commonly identified and psychometrically stable traits from the personality literature, Neuroticism (or Emotional Stability) reflects a dispositional tendency towards experiencing negative affect, such as anxiety and depression, (Cattell &

Scheier, 1961; Widiger, 2009) and is part of the widely used and studied Big Five taxonomy¹⁰ that organizes personality variance into five broad traits (Fiske, 1949; Goldberg, 1981; McCrae & Costa, 1987). Neuroticism is perhaps the personality trait most associated with psychopathology (Kotov et al., 2010; Lamers, Westerhof, Kovács, & Bohlmeijer, 2012; Malouff et al., 2005; Ormel et al., 2013), and is notable as the only Big Five trait that is typically keyed in the maladaptive/negative direction. It is also the broadband personality trait that over time has been the most frequently included as a predictor of interest in conditioning studies (Lonsdorf & Merz, 2017) and was central to early prominent theories as to why some people conditioned more readily or to a greater degree than others (Eysenck, 1970; Pavlov, 1927; J. B. Watson & Rayner, 1920). In some ways, Neuroticism (somewhat aligned with the concept of neurosis, per the Freudian tradition, earlier in the 20th century) was the focal individual difference for human fear conditioning research for most of its early history as an experimental area (H. J. Eysenck, 1962, 1979; Franks, 1956).

In terms of findings from more modern conditioning paradigms (e.g., discrimination learning paradigms), results are mixed, with the majority of studies finding no association between Neuroticism and differential fear conditioning (for review, see Lonsdorf & Merz, 2017). This notably includes studies using physiological measurement (e.g., Hur, Iordan, Berenbaum, & Dolcos, 2016; Martínez et al., 2012; Otto et al., 2007; Pineles, Vogt, & Orr, 2009), behavioral ratings (e.g., Arnaudova, Kryptos,

¹⁰ It should be noted that the Neuroticism factor was found in studies that predate the modern inventories and usage of the term Big Five, starting with seminal work by Thurstone (1934), and that alternative structures to the Big Five also contain Neuroticism factors (Lee & Ashton, 2004; Lee, Ogunfowora, & Ashton, 2005). For the purposes of this dissertation, we will consistently limit ourselves to the Big Five framework when discussing Neuroticism, as well as other traits that comprise the Big Five.

Effting, Kindt, & Beckers, 2017; Lommen, Engelhard, & van den Hout, 2010; Tzschoppe et al., 2014), and/or neuroimaging (e.g., Panitz et al., 2018; Tzschoppe et al., 2014) and also includes studies using relatively large samples (e.g., $N = 217$; Pineles, Vogt, & Orr, 2009). The few studies finding a positive association between Neuroticism and fear acquisition typically find complex effects that require substantial inference and qualification, such as a finding related to a specific subscale (facet) of Neuroticism (Pineles et al., 2009) or a similar construct (Gazendam et al., 2014), being conditional on other experimental or quasi-experimental factors (e.g., executive load, Hur, Iordan, Berenbaum, & Dolcos, 2016; presence of dopamine-related polymorphism, Panitz et al., 2018), or found to be a non-significant predictor when controlling for other conceptually similar variables (Sjouwerman et al., 2018).

To our knowledge there is only one study of the Neuroticism trait in the context of a fear generalization paradigm (Arnaudova, Krypotos, et al., 2017). In this study, Arnaudova et al. (2017) did not find an effect of Neuroticism on levels of fear generalization. Of note, this was a relatively small sample ($N = 58$) that was dichotomized into high and low Neuroticism groups, which creates additional concern about interpretations based on these data (Altman & Royston, 2006).

Both the previously cited study by Arnaudova et al. (2017) and one other study that predated the Arnaudova study, used similar methods, and also dichotomized Neuroticism into high/low groups in a relatively small sample ($N = 55$; Lommen, Engelhard, & van den Hout, 2010), provide results related to Neuroticism and generalized instrumental avoidance. Only Lommen et al. (2010) found a significant effect of Neuroticism, with those higher on Neuroticism demonstrated increased generalized

avoidance. This result was not replicated in Arnauadova et al. (2017). Although both of these studies contain the procedural components needed to assess APIC-G, neither report relevant results. Overall, these results seem to offer, at best, a modest endorsement of Neuroticism as a predictor of generalized avoidance.

Extraversion.

Conceptualized as another core personality trait and part of the Big Five (Costa & McCrae, 1980; Quilty, DeYoung, Oakman, & Bagby, 2014), Extraversion refers to a dispositional tendency towards approach and sociability. Although frequently referred to, both in scientific studies and in the media, as a being a social domain, Extraversion can be conceptualized as capturing general approach and reward tendencies and correlates strongly with questionnaires that explicitly measure those constructs (e.g., Carver & White, 1994; DeYoung, Weisberg, Quilty, & Peterson, 2013; Gomez, Cooper, & Gomez, 2005; Quilty, DeYoung, Oakman, & Bagby, 2014; Smits & Boeck, 2006). Lower levels of Extraversion (introversion) are frequently observed as a correlate of internalizing psychopathology (Kotov et al., 2010; Naragon-Gainey, Watson, & Markon, 2009; D. Watson & Naragon-Gainey, 2014; D. Watson, Stasik, Ellickson-Larew, & Stanton, 2015). Extraversion was a particular focus of earlier conditioning theories that revolved around the assertion that those low on Extraversion (i.e., high on introversion) were particularly vulnerable to fear and anxiety and would more readily condition (Davidson, Payne, & Sloane, 1964; H. J. Eysenck, 1979; J. A. Gray, 1972; Kelly & Martin, 1969).

Modern conditioning studies that have tested Extraversion as an individual difference variable of interest are extremely limited, and there are no available fear generalization or avoidance (generalized or otherwise) studies. One differential fear

conditioning study with an explicitly exploratory aim did not find an overall effect of Extraversion on fear acquisition, but instead found a number of effects related to narrow subscales (facets) of Extraversion (Pineles et al., 2009). Specifically, those lower on a “warmth” or “activity” facets showed maintenance of the fear response. Two other studies using a differential fear learning task replicated the overall Extraversion null result; facets were not tested (Martínez et al., 2012; Otto et al., 2007). Although the role of reward and positive affect has been clearly linked to Pavlovian fear conditioning models using experimental manipulations (e.g., Casa, Mena, Ruiz-Salas, Quintero, & Papini, 2018; Pittig & Dehler, 2018; Pittig, Hengen, Bublatzky, & Alpers, 2018), it appears that at present there is very weak to no evidence for an association between Extraversion and fear conditioning (and, adhering to the aphorism that “absence of evidence is not evidence of absence”, it is not possible to confirm this as a “true” null finding at this time; Altman & Bland, 1995). A potentially compelling explanation for this, besides the relative lack of research in the area overall, is that previous studies have not included reward-related variables, most notably approach-avoidance elements, that might “activate” the significant effect of Extraversion in fear conditioning studies (Aupperle et al., 2011; Pittig, Hengen, et al., 2018; Pittig, Treanor, et al., 2018).

Conscientiousness.

The third Big Five trait of relevance to the current study, Conscientiousness refers to the dispositional tendency towards order, discipline, and goal-orientation (Roberts, Chernyshenko, Stark, & Goldberg, 2005; Roberts, Lejuez, Krueger, Richards, & Hill, 2014). Extreme levels of Conscientiousness have been linked to the presence of psychopathology (Hewitt & Flett, 2007; Roberts, Jackson, Burger, & Trautwein, 2009;

Samuel & Widiger, 2011), with high levels of Conscientiousness-related traits and/or perfectionism linked to elevation on an obsessive-compulsive spectrum of pathology (e.g., Carter, Guan, Maples, Williamson, & Miller, 2016; Hopwood, Schade, Krueger, Wright, & Markon, 2013; L. C. Morey et al., 2002; Samuel & Widiger, 2011), although it should also be noted that the OCD diagnosis itself is, somewhat paradoxically, related to lower Conscientiousness (e.g., Kotov, Gamez, Schmidt, & Watson, 2010; Rector, Hood, Richter, & Bagby, 2002; Wu, Clark, & Watson, 2006), which has been posited to reflect that although there is a tendency towards rigidity and orderliness in those with OCD, actual productivity and achievement is severely impacted and therefore reflected in overall Conscientiousness scores.

Like Extraversion, prior work on the trait in relation to fear conditioning is scarce and there are no studies of fear generalization or instrumental avoidance. To our knowledge, the only available and reported¹¹ fear conditioning results related to Conscientiousness are in the same studies that also tested Extraversion using a Big Five measure (Martínez et al., 2012; Pineles et al., 2009). In Pineles et al. (2009), an overall effect of Conscientiousness was not found, but several facets were related to fear conditioning variables. Those higher on “dutifulness” and “order” showed attenuated fear responses, whereas those high on “self-discipline” demonstrated potentiated fear responding. In contrast to the overall null effect of Conscientiousness found in Pineles et al. (2009), the study conducted by Martínez et al. (2012) found trait Conscientiousness was negatively associated with differential fear conditioning; this is consistent with some

¹¹ Otto et al. (2007) report using the full NEO Five Factor Inventory in their study, but only reported Extraversion and Neuroticism results in their manuscript.

of the facet-level findings from Pineles et al. (2009). Taken together, there appears to be preliminary evidence that Conscientiousness buffers against acquisition of conditioned fear, and this is perhaps driven by lower-level facets. Interestingly, a sizable amount of fear conditioning research, including generalization research, has been published on the effect of self-defined “rules” and experimenter-defined instructions as they pertain to conditioning phenomena (e.g., Ahmed & Lovibond, 2018; Boddez, Bennett, van Esch, & Beckers, 2016; Duits et al., 2017; Mertens & De Houwer, 2016; Wong & Lovibond, 2017) and has found different learning effects depending on differences in these two constructs, suggesting the possible differences in Conscientiousness are associated with these effects and therefore the trait has a larger effect on fear conditioning processes than previously thought.

Interim summary of normative broadband and narrowband individual differences.

Overall, the relative scarcity and heterogeneity of research in normative individual differences does not lend itself to strong conclusions. There appears to be some viable signal slightly rising about the noise, as some studies, especially those with larger sample sizes (e.g., $N > 150$; Flores, López, Vervliet, & Cobos, 2018; Martínez et al., 2012; Pineles, Vogt, & Orr, 2009), have found significant associations between these individual differences and fear conditioning variables. Further, there is some evidence of a positive association between negative affect related narrowband traits (trait anxiety, IU) and heightened responding to stimuli with ambiguous or safe signal values (e.g., Chin, Nelson, Jackson, & Hajcak, 2016; Gazendam, Kamphuis, & Kindt, 2013; Haaker et al., 2015; Haddad, Pritchett, Lissek, & Lau, 2012; Kindt & Soeter, 2014; Morriss, Saldarini,

Chapman, Pollard, & van Reekum, 2019), which provides some evidence that these traits are positively related to fear generalization. However, at present it is most reasonable to conclude that the bulk of the research in this area has yet to be conducted. This is especially true for studies of instrumental avoidance, which at this point in time consists mainly of two studies using largely identical paradigms with results that contradict each other. Given the current state of the literature, it appears investigators are still mainly in the “exploratory” phase, as opposed to the “confirmatory” phase of the scientific research process (Jebb, Parrigon, & Woo, 2017; van’t Veer & Giner-Sorolla, 2016), which is also a sentiment that is discussed by Pineles et al. (2009) and contributes their conclusion that the personality and conditioning relationship is a complex one that still requires considerable additional research. This is particularly true for traits that are not conceptualized as primarily reflecting negative affect (e.g., Conscientiousness and Extraversion) that can be reasonably expected to influence approach-avoidance conflict, and therefore generalization processes during these conflicts, but have not been studied at all. We also contend that this represents a myopic view in which investigators are attempting to conduct studies with negative affect traits that have face-valid relevance to internalizing psychopathology, yet ignoring variables that are certainly related to internalizing psychopathology but are not as obvious a target of study as the negative affect variables.

As a next step towards confirmatory hypothesis testing and a more consistent approach to examining normative individual differences testing in fear conditioning, we support the recommendations made by Lonsdorf and Merz (2017) in regards to standardization of methodology (also see Beckers, Krypotos, Boddez, Effting, & Kindt,

2013; Lissek, Pine, & Grillon, 2006; Lonsdorf et al., 2017). The following four factors appear to be the most important inconsistencies or problematic decisions to address: 1) transforming inherently dimensional data into categorical creates numerous problems, including those related to statistical inference (Altman & Royston, 2006) and inconsistencies in cut-scores and group ranges (Lonsdorf & Merz, 2017); 2) relatively small sample sizes that are underpowered for detecting the relatively modest effects that are found in psychological science research (Lattin, Carroll, & Green, 2003; Lykken, 1968; Paunonen & Ashton, 2001), especially those relationship measured across different measurement methods (e.g., Campbell & Fiske, 1959; Kozak & Miller, 1982; Lang, Levin, Miller, & Kozak, 1983), and potentially lead to overinterpretation of unstable, sample-specific effects (Bakker, van Dijk, & Wicherts, 2012; Fiedler, 2011); 3) continued usage of univariate approaches (i.e., only using one individual difference as a predictor or conducting separate models for each predictor without also including an omnibus multivariate test) that do not take advantage of multivariate statistical techniques to sharpen or control for interacting effects between individual differences measures (Lonsdorf & Merz, 2017; Sjouwerman et al., 2018); and relatedly, 4) inconsistent usage of different individual difference measures and, with two exceptions (Gazendam et al., 2014; Pineles et al., 2009), an apparent lack of empirical interest in applying measures that capture hierarchical variance structures which capture individual differences of interest at different levels (e.g., personality traits and facets) and can provide insight into broad vs. specific individual difference influences on fear conditioning processes within a single study. The lack of sophisticated individual difference measurement and corresponding advanced statistical methodology is particularly notable and likely

problematic given copious research showing the substantial advantages in using hierarchical and empirically-derived multidimensional personality systems to quantify human variance, amongst them improved prediction of behavior and stronger links to biological substrates and meaningful clinical correlates (e.g., Corr & Matthews, 2009; DeYoung et al., 2010; DeYoung, Quilty, & Peterson, 2007; Goldberg, 1999; Kotov et al., 2017; Paunonen, 1998; Paunonen & Ashton, 2001). Additionally, lack of standardization regarding experimental parameters and quantification of experimental outcome variables (e.g., differences in how startle EMG is quantified can lead to different predictive validity, statistical testing power, and/or interpretations; (Bach et al., 2018; Bradford, Kaye, & Curtin, 2014) might contribute to inconsistent replication and increased heterogeneity of findings that hinder synthesis in the area.

Individual difference variables with potential relations to generalization variables.

Beyond the individual difference variables discussed previously, there are a number of candidate individual differences that, to this point, have not been tested as predictors of fear or avoidance conditioning. These variables represent potentially meaningful sources of variation in APIC and its constituent processes, either due to measuring relatively unstudied sources of variation (e.g., pathological manifestations of personality variables) or measuring established individual differences variables with different techniques that result in more refined or nuanced data. Candidate variables that are relevant to the current study are discussed below, with a focus on establishing what is unique about these variables compared to those currently employed in generalization studies.

Dimensional pathological individual differences.

There is considerable evidence that normative personality measures also capture pathological variance (e.g., Mahaffey, Watson, Clark, & Kotov, 2016; Watson & Naragon-Gainey, 2014), and preliminary evidence suggests that these measures are sensitive enough to link certain normative traits with conditioning abnormalities (e.g., Gazendam et al., 2014; Martínez et al., 2012; Pineles, Vogt, & Orr, 2009). However, measures that specifically target pathological variance (i.e., the tails of the normal distribution assessed by the normative personality measures) offer multiple benefits over normative measures, including broadening the range of the predictor being tested (which likely reduces potential ceiling effects in the normative personality measures, with the likely exception of Neuroticism; e.g., Watson & Naragon-Gainey, 2014) and incorporation of more unusual or lower base-rate behaviors and internal processes that are relatively rare in the general or even clinical population (e.g., moral/sexual/religious obsessions; Fullana et al., 2010; Haslam, Williams, Kyrios, McKay, & Taylor, 2005; or visual hallucinations; e.g., Johns & van Os, 2001; Stefanis et al., 2002). That said, the most commonly used measures of pathological personality (e.g., MMPI, PAI) might have been designed to be somewhat too attuned to clinical problems, or contain too much content that is keyed to specific categorical clinical disorders, for ideal use outside of clinical populations and perhaps have a floor effect (for discussion of this issue, see Costa & McCrae, 1992; Krueger & Markon, 2014; Samuel, Simms, Clark, Livesley, & Widiger, 2010). This suggests there is a need for measures that capture a higher upper limit than normative measures while maintaining sensitivity to the lower limits of the range and ensuring that low base-rate symptoms are not overrepresented in item content

(i.e., obtaining a distribution that, although skewed, maintains substantial density towards the lower end of the trait). Fortunately, this type of measure aligns closely with a (relatively) recently invigorated push towards dimensional classification of psychopathology (e.g., Hopwood et al., 2018; Kotov et al., 2017). Currently used instruments of this type have yielded important candidate variables for fear and avoidance generalization studies, including higher order internalizing traits such as Negative Affectivity/Negative Emotionality/Emotional Dysregulation with strong fear, anxiety, and distress components, as well as other traits that might be related to increased and potentially maladaptive approach under threat, such as Disinhibition/(Dis)constraint/(low) Inhibitedness (Krueger, Derringer, Markon, Watson, & Skodol, 2012; Kushner, Quilty, Tackett, & Bagby, 2011; Simms et al., 2011). The advantages of using these types of instruments to measure these dimensions is that they can be used in an assumed normative population (e.g., undergraduates, general community members) without extensive concern regarding range-restriction or zero-inflation. It should be also noted that although the cited measures and studies mainly reflect efforts to measure personality disorders using a dimensional approach, these measures and the conclusions from studies employing them are quite valid when applied to other forms of psychopathology (as well as when considering the issue from the point of view that there is no viable distinction between personality disorders and other forms of psychopathology (L. A. Clark, 2005; Kotov et al., 2017; Krueger & Markon, 2014; Widiger & Shea, 1991). Overall, these types of variables show great promise for APIC investigations given the breadth of pathological variance measured and the interest in pathological manifestations of APIC from an etiology and intervention standpoint and

can be conceptualized as extensions of the normative variables that are already employed in the area.

Lower-level personality variables.

Part of the appeal of applying personality methods to the study of fear conditioning phenomena is the broad coverage of variance and content without creating an unrealistic burden for participants, as well as potential reliability and validity issues, through administration of many narrowband measures for a variety of disparate content areas. A potential drawback of using personality methods is if measurement is only conducted at the trait level of personality, which is perhaps too far removed from the fine-grained psychological and neurobiological substrates of experimentally induced fear conditioning processes¹². A viable solution to this issue is to consider that personality structure is naturally hierarchical (Costa & McCrae, 1995; DeYoung, 2006; McCrae et al., 2008) and to use measures that provide viable measurements at different levels of the hierarchy. If considering the “trait” level (e.g., Big Five traits) as the relative anchor or center that we then use to define other levels of the hierarchy, then those representing a broader superordinate set of factors are “higher-level” (DeYoung, 2006; Digman, 1997) and, most importantly for this endeavor, those more narrow factors that are subordinate to the traits are “lower-level” (Costa & McCrae, 1995; Goldberg, 1999).

Lower-level personality variables can vary in degree of specificity (i.e., how far down the hierarchy they are in relation to traits), with some of the most common narrowband measures that focus on symptoms of specific disorders perhaps representing

¹² The author is making a general statement that and is not explicitly referring to or intending to remind the reader of the person-situation debate (e.g., Mischel, 1977), nor implying that personality psychology is not applicable to experimental (situational) psychology.

one of the lowest levels of personality (Clara, Cox, & Enns, 2003; Kotov et al., 2017; Markon et al., 2005; Widiger et al., 2019). In terms of commonly studied lower-levels of personality, the most relevant ones here are aspects (DeYoung et al., 2007) and facets (Goldberg, 1999), with facets representing a specific type of content that is part of a trait (e.g., as mentioned previously, Warmth is a facet of Extraversion on some scales), and aspects are below traits and contain a subset of facets within them. These lower-level variables might be more refined predictors of fear conditioning processes than the traits, especially considering that traits include aspects or facets that might interact with other aspects or traits within the same trait, and that these interactions might be more informative than the individual univariate relationships. For example, Neuroticism is frequently identified as having both strong anxious and depressive elements (Barlow, Ellard, Sauer-Zavala, Bullis, & Carl, 2014; Barlow et al., 2013; L. A. Clark & Watson, 1991; D. Watson, 2005). In some cases, these parts of Neuroticism work synergistically to produce an effect (e.g., increased substance use and poorer health outcomes; e.g., Burns & Teesson, 2002; Mykletun, Overland, Aarø, Liabø, & Stewart, 2008; Sartorius, Üstün, Lecrubier, & Wittchen, 1996), whereas in other cases they differentially relate to a criterion (e.g., mortality, positive affect, or approach behavior; L. A. Clark & Watson, 1991; Mykletun et al., 2009; Shankman & Klein, 2003). When modeled separately from Neuroticism in a multivariate model, these effects can be clarified, but when modeling only the Neuroticism trait it is impossible to disentangle the distinct contributions of anxiety and depression. This is particularly notable for fear conditioning given that anxiety is consistently linked with fear conditioning abnormalities, whereas depression is not (e.g., Dibbets, Broek, & Evers, 2015; Jovanovic et al., 2010), and there is evidence of

a “dampening” effect in depression physiology (e.g., Bylsma, Morris, & Rottenberg, 2008; Sloan & Sandt, 2010) that can attenuate the signal in common fear conditioning measures, such as startle EMG or skin conductance. Thus, if Neuroticism is assumed as a marker of anxiety and entered in a model, but depression is not modeled (“covaried out”), false negatives (type II error) are more likely because the depression variance within the trait might “wash out” the anxiety signal.

Even when including multiple lower-level variables instead of a single higher-level trait there are potential statistical effects that affect interpretation and potentially offer added explanatory power. For example, DeYoung and colleagues (2007) established that two aspects (lower-level factors located between traits and facets in a hierarchical model) of Conscientiousness, termed Industriousness (tendency towards productivity, focus, and goal-oriented) and Orderliness (tendency towards control and rule-following), are differentially associated with Neuroticism when calculating zero-order correlations: Industriousness is significantly anticorrelated with Neuroticism, and Orderliness has almost zero correlation. However, when including both in the same model predicting Neuroticism, Orderliness was now significantly correlated with Neuroticism. This is an example of a suppression effect, in which two positively-related variables are differently related to a third variable, but these relationships are potentially suppressed in simple bivariate procedures (MacKinnon, Krull, & Lockwood, 2000; Paulhus, Robins, Trzesniewski, & Tracy, 2004). Put another way, the shared source of variance between the two predictors needs to be controlled for before the unique relationships for each predictor emerge.

Again, the use of lower-order personality traits is an attractive compromise for fear conditioning studies, as they balance resolution with generalizability. They might be particularly well-suited for studies of fear and avoidance generalization because these processes are naturally more complex than basic discrimination learning and therefore use more complex laboratory tasks with an increased number or depth of experimental parameters, and the lower-level traits can be more easily operationalized in terms of behavior and adaptations to specific situations (e.g., DeYoung, 2015; DeYoung & Krueger, 2018; Jackson et al., 2010). Further, it might be a de facto necessity to use lower-order traits in this type of research if aiming to inform future research on anxiety- and trauma-related pathology, as the higher-order traits (e.g., Neuroticism) lump together variance that needs to be split out (e.g., anxiety, depression) to usefully inform the development of etiological models and treatments for this pathology.

Trait fear as a complementary construct to trait anxiety.

As noted previously, trait anxiety¹³ has dominated the individual differences work in the fear conditioning field, and results are mixed. A possible explanation lies in the assertion that the trait anxiety construct is too broad to be useful and that measures of trait anxiety are prone to capturing depression-related variance: the strength of the trait anxiety signal is not strong enough to overcome the noise inherent to the measure, particularly when it is the anxiety-specific variance that is the basis for hypotheses regarding fear conditioning constructs. To further complicate the issue, even the term

¹³ We acknowledge that “trait anxiety” is purportedly measured by different instruments and therefore operationalized slightly differently depending on the measure used. We continue to use the previously stated definition of “disposition towards experiencing anxiety in the moment” that is most commonly associated with Spielberger and colleagues’ (1983) definition and associated questionnaire, the STAI.

“anxiety” is problematic in this context. There is considerable variability in terms of how people, both lay and scientific expert, define anxiety (trait or not) and the related construct of fear (Joseph E. LeDoux, 2017; N. McNaughton, 2011; Perusini & Fanselow, 2015; Shackman & Fox, 2016; D. Watson, Stanton, & Clark, 2017), with some differentiating between the two terms and some using them interchangeably to the point that they are synonyms for each other (Gaylin, 1979; N. McNaughton, 2011). This problematically affects both how participants respond to trait anxiety questions (with items on the same scale that indiscriminately include the words fear, anxiety, worry, stress, terror, etc. and might be prone to significant error due to idiographic interpretations of those words) and how researchers conceptualize, implement, and interpret their studies of trait anxiety. In short, trait anxiety to some might not be trait anxiety to all.

Taken together, trait anxiety (as conceptualized in the majority of current empirical research) is rife with construct validity issues (as classically defined by Cronbach & Meehl, 1955) and does not appear to be a useful construct as it pertains to conditioning studies. That said, regardless of how they are defined (or if defined as a single construct), there is clearly a need to accurately and usefully measure what people commonly refer to as anxiety and fear using a transdiagnostic, dimensional approach – it just needs to be a measurement that does not hinge on continued reification of a flawed trait anxiety construct. Thus, an alternative or modification to the trait anxiety construct and its measurement is needed.

One proposal is to subdivide trait anxiety into empirically-supported factors that map more closely to the specific experiences referred to as anxiety, such as somatic vs

cognitive experiences (Grös, Antony, Simms, & McCabe, 2007; Ree, French, MacLeod, & Locke, 2008). Another proposal is to use an undifferentiated term of “anxiety and fear” when referring to the general concept and for more precise discussion to move towards taking a more granular view of these constructs and when possible define them in terms of the specific threat processes and conditions (Shackman et al., 2016; Shackman & Fox, 2016). Finally, another proposal has been to formally distinguish between trait anxiety and trait fear by clarifying the behavioral, neurobiological, and clinical properties of each and to use these as referents for future operationalization (Sylvers, Lilienfeld, & LaPrairie, 2011). The proposed solutions are not necessarily incompatible with each other and all three likely represent important steps forward for the field – however, for the purposes of this dissertation, we continue discussion with a focus on the third solution of separately measuring trait fear.

Trait fear, as measured separately from trait anxiety, is generally defined as a disposition to react defensively and experience subjective distress in response to an acute threat (Perkins, Cooper, Abdelall, Smillie, & Corr, 2010; Sylvers, Lilienfeld, & LaPrairie, 2011); this also helps define trait anxiety by defining it as reaction to an uncertain or distal threat. It should also be noted that there are differences in how scientists specifically define and operationalize trait fear, especially in regards to what is defined as a *lack* of fear (e.g., Kramer et al., 2019; Sylvers, Lilienfeld, & LaPrairie, 2011) or how approach behavior relates to the construct (N. McNaughton, 2011; Perkins et al., 2010), but that the provided definition represents a conceptual overlap. Most importantly, the trait fear construct appears to capture variance that is conceptually related to and might predict fear conditioning processes: fear conditioning paradigms using cue-based

manipulations (e.g., a CS+ or CS- visual stimulus) create acute threat conditions that are proposed as the situation in which trait fear differences emerge. Further, those higher on trait fear would be expected to show greater fear generalization, as trait fear represents a probabilistic tendency to react fearfully when controlling for situation (i.e., lower threshold), and therefore those higher on the trait would have a lower threshold for reacting to GSs. Finally, trait fear is conceptualized as not only an index of internal responsivity but also of behavioral tendency (e.g., escape; Perkins, Cooper, Abdelall, Smillie, & Corr, 2010) and would be expected to correlate with avoidance tendencies in instrumental conditioning paradigms. Based on these assumptions, it follows that trait fear is a dimension with great appeal for APIC research, and might be a “cleaner” signal of the negative affect that underlies generalized fear and avoidance than trait anxiety or a higher order trait such as Neuroticism.

Methodology of Individual Differences

Conceptual rationale for a dimensional individual differences approach.

Drawing from our conclusions from the previous sections, it is clear that more a refined individual differences approach is necessary to push forward fear conditioning science. As previously stated, there are many benefits to this approach. Instead of rehashing this point, we turn briefly to a conceptual rationale for incorporating dimensional individual differences-based approaches into fear conditioning work, and generalization work in particular. This rationale is based on 1) current prominent frameworks in the psychopathology field and 2) historical calls for the integration of the individual differences and experimental psychology subfields.

The relatively recent introduction of the Research Domain Criteria (RDoC) by the National Institute of Mental Health (RDoC; Insel et al., 2010) has shifted research priorities towards a series of conceptually-linked systems (e.g., Negative Valence) that contain multiple constructs (e.g., Acute Threat, Potential Threat) measured across multiple units of analysis (e.g., circuits, physiology, behavior), with an emphasis on neurobiological or neurobiologically-linked measures (National Institute of Mental Health, 2016), and away from the DSM framework that (problematically) classifies psychopathology using expert-determined categorical categories (L. A. Clark, Cuthbert, Lewis-Fernández, Narrow, & Reed, 2017; Krueger, Hopwood, Wright, & Markon, 2014). This focus nominally benefits fear conditioning research given the focus on brain-based models of Pavlovian and instrumental conditioning (H. Kim, Shimojo, & O'Doherty, 2006; J. J. Kim & Jung, 2006; Maren, 2001). However, linking the RDoC constructs to meaningful measures of variation in human behavior, particularly pathological variation, will be difficult without improved individual differences methodology (Krueger et al., 2014; Lilienfeld, 2014; Lonsdorf & Merz, 2017; Patrick et al., 2013a; Shackman et al., 2016; Shackman & Fox, 2018; Shackman & Wager, 2019).

Since the early 1950s, scientists have identified a tradition of “two disciplines” in psychological science research, in which the “correlation scientist” is concerned with the covariation of a wide number of descriptive variables in a large sample and discounts within-subject processes, and the “experimental scientist” views these variables as “error” in their goal of carefully controlling an experiment (while habitually underestimating the difficulty of true experimental control; Lykken, 1991) to observe a differences between conditions or groups that provide insights on a specific mechanism

(Cronbach, 1957, 1975). This conflict has been noted in more recent reviews and studies as hindering scientific progress, (e.g., Patrick et al., 2013; Patrick & Hajcak, 2016; Yancey, Venables, & Patrick, 2016), including those from fear conditioning researchers (Gazendam et al., 2014; Lonsdorf & Merz, 2017). with the most notable effect of this conflict being incompatible or ambiguous interpretations for groups of studies that purportedly measure the same constructs but use very different approaches (e.g., studies of the anhedonia component of PTSD have very different results depending on if they are from the individual differences, (e.g., Armour et al., 2015), or experimental tradition (Nawijn et al., 2015).

Statistical rationale for a dimensional individual differences approach.

One of, if not the greatest, barriers to wide-spread implementation of a dimensional individual differences approach to the study of fear conditioning is that it would require a shift to new statistical methodology. We start with identifying what makes the current statistical paradigm in the field unsuitable to individual difference. At present, the workhorse analytic technique for the fear conditioning field (and a good percentage of experimental studies in psychology overall) is analysis of variance (ANOVA) and techniques from the ANOVA family (ANCOVA, MANCOVA), particularly repeated-measures ANOVA (rmANOVA). As the standard ANOVA framework is fundamentally a test of mean differences, it is a natural fit for experimental work and factorial structures. Repeated-measures ANOVA is a particularly good fit for conditioning work because it allows testing of differences in both within-subject and between-subject variances; many fear conditioning questions revolve around factorial within-subject manipulations related to the signal value of a particular stimulus that is

presented multiple times (e.g., “is response to CS+ > CS-? For all participants”).

Unfortunately, rmANOVA is subtly, but decisively, unsuited to a comprehensive study of fear conditioning variables as they related to dimensional individual differences, including personality variables and dimensional models of psychopathology that continue to gain traction in the RDoC era (e.g., Conway et al., 2019; Hopwood et al., 2018; Kotov et al., 2017). A primary reason for this is that an overwhelming amount of evidence establishes personality variables as dimensional (Corr & Matthews, 2009; Trull & Durrett, 2005) and that artificially categorizing them for primary analyses is usually not defensible, regardless of technique or cut-score used (Altman & Royston, 2006; Dawson & Weiss, 2012; Iacobucci, Posavac, Kardes, Schneider, & Popovich, 2015; Irwin & McClelland, 2003; MacCallum, Zhang, Preacher, & Rucker, 2002; Pittenger, 2004). Continuous predictors are not possible in the rmANOVA framework¹⁴. Even if artificially dichotomized variables are used as “preliminary” predictors within an rmANOVA, the results are substantially unstable and biased, leading to the possibility of drastically different results between dichotomized and continuous variable predictors (Altman & Royston, 2006; J. M. Taylor & Yu, 2002). Clearly, proper dimensional individual differences work in the fear conditioning area cannot be conducted using rmANOVA, and more recent investigations that have used more appropriate techniques are responsible for some of the more comprehensive and robust recent findings in the field (e.g., Gazendam et al., 2013).

Issues with rmANOVA applied to generalization data.

¹⁴ Although continuous covariates can be modeled with ANCOVA based techniques, these are limited in that they only help account for between-subjects variation in the context of a categorical predictor (i.e., adjust the group means by the covariate) and cannot function as true predictors of the outcome variable.

Moving to the area of fear generalization in particular, we encounter further issues with the rmANOVA framework. Vanbrabant and colleagues (2015) provide a detailed critique of the rmANOVA approach for fear generalization studies (which applies equally to avoidance generalization studies) and compare it in a proof-of-concept investigation to multilevel modeling techniques (also referred to as mixed models or hierarchical modeling). We briefly summarize the critique from Vanbrabant et al. (2015) and then review why multilevel modeling (MLM) is a superior technique for generalization data, including data from the current study.

Vanbrabant et al. (2015) contend that rmANOVA is inappropriate for generalization data due to 1) the previously discussed issue of needing to dichotomize continuous predictors for use in an ANOVA; 2) the generalization gradient is, by definition in the ANOVA framework, treated as a categorical variable; and 3) the design of generalization experiments inherently violate rmANOVA assumptions and therefore bias statistical tests to an unacceptable degree. Regarding categorical generalization gradient, this is perhaps the only tenuous or debatable claim made by Vanbrabant et al. (2015). It is likely true that representing the generalization stimulus continuum as dimensional better captures generalization as it exists outside of the lab (e.g., Sims, 2018; Tenenbaum & Griffiths, 2001), but it is not clear yet if the experimental instantiations of generalization stimuli for human participants are sophisticated enough to capture a truly continuous gradient or if that would be a meaningful endeavor. That said, of two imperfect options it appears that representing generalization stimuli as a continuous variable is the preferable option from a statistical standpoint. In terms of violations of sphericity, this is the area in which generalization research creates the most conflict with the rmANOVA framework.

In experimental generalization research there is, by definition, a violation of sphericity (i.e., variance of differences between the paired combinations of repeated-measure factors are unequal). This is clearly seen when considering that the variance of differences between a CS- and its closest approximation will be much smaller than the variance of the difference between the CS+ and CS-, which typically have the least amount of variance due to the conditioning acquisition procedure (for expansion of this example and explanation see Vanbrabant et al., 2015). Corrections for violation exist, but they are limited and potentially result in a higher chance of a type II error, especially compared with other techniques for repeated-measures data (e.g., Misangyi, LePine, Algina, & Goeddeke, 2006; Vanbrabant et al., 2015). An additional point made by Vanbrabant et al. (2015) is that all cases subjected to rmANOVA must be complete (i.e., can't have any missing data), which results in loss of power through participant exclusion and, in some cases, uneven sample sizes (and between-subjects different degrees of freedom) for different analyses from the same study if different dependent variables are analyzed with separate analyses (e.g., missing psychophysiological data for 5 participants compared with only missing expectancy rating data for 1 participant will result in different sample sizes for each analysis unless viable risk rating data is excluded). It should also be noted that the specific points made by Vanbrabant et al. (2015) are also made by other authors commenting on issues in the general area of experimental psychophysiological research (Bagiella, Sloan, & Heitjan, 2000; Kristjansson, Kircher, & Webb, 2007; Vasey & Thayer, 1987), psychiatry (Gueorguieva & Krystal, 2004), and other psychological and medical fields of study (Hoffman & Rovine, 2007; C. Krueger & Tian, 2004).

Using multilevel models with generalization data.

Vanbrabant et al. (2015) and others (e.g., Hoffman & Rovine, 2007; Kristjansson, Kircher, & Webb, 2007) propose MLM as the solution to the problems inherent rmANOVA. MLM can directly address the specific issues identified in the rmANOVA framework: it can model continuous predictors (both between-subjects and within-subjects), has less stringent assumptions than rmANOVA (most notably it does not require sphericity), can exclude missing data pair-wise, and generally contains all the advantages of the regression framework while allowing for the dependencies that are seen in the experimental variables that are modeled in repeated-measures designs (Gelman & Hill, 2006). The MLM framework is also more robust for hypothesis testing and to multiple comparisons issues due to technique used to pool variance for estimates of population variance used in hypothesis testing (Gelman, Hill, & Yajima, 2012). The only relative drawback is that MLM typically needs larger sample sizes than those used for rmANOVA (Maas & Hox, 2005), however, this is mitigated both by the option of using fewer parameters in MLM if 1) obtaining a proper sample size is an issue and 2) the reduction of model complexity can be justified (McNeish & Stapleton, 2016; McNeish & Stapleton, 2016). The requirement for relatively larger sample sizes is particularly important if using MLMs in the context of dimensional individual differences research, as a larger sample size will be required to sufficiently detect stable effects related to those differences (e.g., Kotov, Gamez, Schmidt, & Watson, 2010; Roberts, Walton, & Viechtbauer, 2006; Schönbrodt & Perugini, 2013).

Stepping back from the specific details of MLM, we also contend that there is a conceptual difference between the rmANOVA compared with the MLM approach that

summarizes the primary rationale for generalization research to fully embrace MLM and other more advanced techniques. The rmANOVA framework is predicated on the viability of using a measure of central tendency (marginal means in most cases, specifically group means and repeated-measures factor means) to answer a scientific question that revolves around if two or more categorical variables are significantly different. In this case, variance around the means are reduced to a metric of uncertainty (i.e., error) that influences hypothesis tests (e.g., if there is a large group difference but also substantial variability around the mean, then the test will likely not be significant). MLM, on the other hand, allows for modeling and testing of these means (termed fixed effects, equivalent to the effects modeled in standard multiple regression), but also allows for the modeling of individual-level variance (termed random effects). Put another way, MLM reclaims what the rmANOVA treats as error and uses it as part of the model. This flexibility and the move away from only testing mean differences is the most compelling argument for using the MLM framework over rmANOVA. Unless an investigator can be sure that 1) members of a group are more similar to each other than to members of another group, 2) the mean response when holding all predictors constant (i.e., the intercept) does not meaningfully differ across participants, or 3) that participants do not meaningfully differ in their pattern (i.e., slope) of responses across one or more repeated-measures, rmANOVA is likely not a good fit for their data and can lead to biased estimates and interpretations. For a worked example demonstrating these points, see Vanbrabant et al. (2015). Within the context of the current dissertation, we use another modification of the previously established example of the Rottweiler and the child to broadly illustrate these points. Let us establish that, in addition to the child from the

original example, that there are ninety-nine other children who live on his block (it's a big block) who have encountered the Rottweiler (CS+), the toy poodle (CS-), and the other dogs that resemble the Rottweiler to some degree but are not dangerous (GS). This gives us a sample size of 100 children that have generalized fear from the Rottweiler to the similar looking dogs. If we took the average response (an arbitrary unit of fear) to the continuum of dogs, we would likely end up with a reasonable approximation of a fear generalization gradient, with the highest response to the CS+, and responses precipitously diminishing from the CS+, with perhaps a slight "bump" or more shallow decline to the closest approximation of the CS+. Further, based on what we know about APIC-G, we might expect that this fear generalization gradient might help us predict future generalized avoidance for this block, or even for future blocks in similar neighborhoods. However, for the parent or parents of each individual child, this might not be so reassuring – they are concerned about their children's level of fear and would like to be able to intervene in an effective manner. One option is to allow the children to go out and encounter the CS+ and GSs so many times that the parents have enough data to make a reasonable prediction for themselves (i.e., we have sampled one participant many times), but the parents might still not be sure based on this information how their child would react to a novel, but similar, dog. Another option is to leverage the collected data to predict on the individual level. Our first step away from prediction entirely on a group mean is to see if we could determine "baseline" level of fear for each child, as it stands to reason that some are very fearful and some are temperamentally less fearful. We could then use this information to modify the prediction based on the overall group mean. But this does not actually tell us anything about differences in generalization – two children

with the same baseline fear could generalize very differently. At this point, we would then need to determine the pattern of responding to each dog for each child. This would provide us with a more accurate prediction for each individual child.

It's at this point that we pause and state that from the experimental point of view, we have two issues now: 1) individual variation over time for our arbitrary unit of fear might be stronger than the stimulus-specific variation (i.e., more noise than signal) and 2) if we individually predict a child's generalization gradient we lose our ability to make inferences about similar children. The first issue is what ultimately makes it difficult to produce individual-level prediction in experimental work, which typically uses "noisy" methods, but it's the second issue that is addressed through MLM and not rmANOVA. In rmANOVA, it is true that if we model each child as their own group of $N = 1$, we have completely lost any ability to infer what similarities amongst children account for similar generalization gradients, both in terms of dispositional predictors (e.g., personality) and in terms of which dogs (stimuli) are "driving" differences in generalization. If we take the opposite approach and create two groups of $N = 50$ based only on a single dispositional variable, we are now not sure if this represents a "true" difference or if it's an artifact of how we created our groups. That leaves us with treating these children as a single group of $N=100$ individuals, which if done using rmANOVA means we no longer have any form of between-subjects factor (i.e., we do not have two or more means to test) and cannot predict how generalization varies based on dispositional traits. In the MLM framework, we do not have this issue, as we can measure how changes in a dimensional dispositional trait (i.e., between-subjects variability) are associated with differences in a

generalization gradient (i.e., within-subjects variability). In the rmANOVA framework, this would have been impossible.

Finally, we note that although the majority of generalization studies, including the vast majority of the studies cited previously in this document, use rmANOVA, they still provide important contributions to the literature and are likely true-positive effects. The issue, in our view, is not whether they are false-positives or not. The issues are that effect size estimates are likely biased, and conclusions based on these effects are relatively imprecise and therefore imperfect foundations for dimensional individual differences work.

Summary and Conclusions

At present, the fear and avoidance conditioning literature represents a concerted, systematic effort to increase understanding of the mechanisms that underlie fundamental dimensions of human behavior – the need to learn and predict danger, discriminate between threat and safety, and escape or approach as appropriate for a given context. These processes and their behavioral consequences can be adaptive or maladaptive, depending on the context.

Aberrations in these processes that result in maladaptive generalization, as well as their covariation (i.e., generalized fear leading to generalized avoidance), are proposed as core mechanisms of anxiety and trauma psychopathology. When using a categorical approach (e.g., DSM disorders), results from empirical research and observations from clinical endeavors coalesce into support for this proposal. In particular, we have strong support for the role of Pavlovian (emotional-passive) forms of generalization as a pathological correlate. Direct empirical support for instrumental avoidance (active-

behavioral) generalization, as well as its relation to Pavlovian generalization, is lacking, although clinical observations and basic conditioning theories suggest it is also a pathological correlate. Empirical work in this area is just beginning, and it will likely be years until sufficient literature exists to support broad conclusions in this area. That said, enthusiasm for continued research using categorical DSM disorders is dampened by established concerns with the construct validity of these disorders and that conditioning studies of categorical disorders typically use suboptimal sample-sizes and statistical techniques.

In parallel to investigations on generalization and pathology, there has been empirical interest in identifying normative correlates of fear and avoidance generalization, which can inform our understanding of these processes as they pertain to the majority of the population while providing insight into how normative traits and behaviors can convert into pathological and maladaptive forms. Initial studies of fear generalization have yielded mixed results, with findings frequently contradicting one another and notable inconsistencies evident in methodological and statistical approaches. Studies of instrumental avoidance generalization are even less interpretable. Additionally, the individual differences traits that have been the focus of previous research are relatively narrow in terms of content area and likely do not capture enough sources of variance to provide optimal predictive validity. There is also concern regarding the resolution of the traits tested in previous studies, with almost every reviewed study restricting analysis to a single level of the established hierarchical structure of personality and ignoring a “middle level” of personality variance that might represent an optimal balance of content specificity and sensitivity.

When considering these two lines of related research together, we see that each has a unique limitation and then that both also share another limitation. For studies of patient groups, there is concern that the grouping variables used (DSM-diagnosed disorders) do not reflect a coherent system of psychological or neurobiological processes, and therefore the strength of the relationship between these groups and more fundamentally and coherent Pavlovian and instrumental processes, is artificially reduced. Put another way, the heterogeneity inherent to categorical disorders limits the ability to make strong conclusions and, potentially, hinders translational efforts. For studies using normative personality traits, a narrow focus on negative affect related traits, as well as few researchers leveraging the robust literature establishing the utility of a hierarchical approach to improve prediction, has resulted in an incomplete picture of the relations between individual difference traits and generalization processes. For both lines of research, choices in regards to needed sample size, quantifying individual differences (e.g., study-specific cut scores for dimensional traits) and statistical approach (e.g., rmANOVA for designs that by default violate sphericity and are inappropriate for dimensional predictors, resulting in a focus on means at the expense of predicting variances) represent major limitations and, given their entrenchment in the field, likely impediments in needed future progress. Our overarching conclusion is that despite a number of promising results in the area, key relationships are still either unclear or untested, and both represent major gaps in the literature. In terms of gaps in the literature related to a lack of clarity in extant research, we highlight the following for their relevance to the current study:

1. We do not have clarity regarding how the negative affect traits (e.g., Neuroticism, trait anxiety, trait fear), especially when measured and analyzed in their naturally dimensional form, relate to fear or avoidance generalization. Similarly, it is not clear how these traits relate to the covariation of generalized fear and avoidance.
2. It is still extremely unclear how variables that are not directly reflective of negative affect (e.g., Conscientiousness, Extraversion) relate to fear or avoidance generalization, despite their relevance to important dimensions of human behavior and identified pathological processes. As with the negative affect variables, it is also not clear how these variables relate to the covariation of generalized fear and avoidance; particularly in the context of approach-avoidance conflicts that allow classification of adaptive vs maladaptive outcomes and might be better predicted by the variables such as Extraversion or Conscientiousness than the negative affect variables.
3. Only one study so far (Pineles et al., 2009) has compared the predictive power of higher-level personality factors (traits) to lower-level factors (facets) in the context of fear conditioning work, and this was a largely exploratory endeavor. Further, the relatively atheoretical nature of that study resulted in limited analysis and interpretation of the lower-level factors.
4. Although studies of psychiatric patient populations have been fruitful, it is not yet clear if the identified pathological markers represent unique substrates of categorically defined disorders, or if these markers significantly covary with the personality dimensions that have been supported as a superior etiologic and taxonomic model of psychopathology (i.e., better “carves nature at its joints”).

There are also still questions that have not been addressed yet in the empirical literature and represent important next steps. We propose that the following gaps represent the most pressing next steps for the area:

1. In general, there is a lack of research utilizing experimental approaches to instrumental generalization. Most notably, there are no available studies of pathological individual differences (measured categorically or dimensionally) in instrumental generalization, nor studies that take a hierarchical approach to studying individual differences and examine higher vs lower-level variables in this sub-area.
2. There are no available studies of individual differences in APIC-G (and, by extension, APIC-CS+). Notably, this means the field has not yet established if individual differences can improve predication of the APIC processes (which is particularly important given the relatively weak predictive power of Pavlovian processes), nor has it identified which individual differences moderate the relationship between Pavlovian and instrumental responding. This, of course, means we do not have empirical evidence of what individual differences are risk or protective factors for APIC processes.
3. There is evidence that putatively normative measures of personality (e.g., Big Five) predict pathological processes, behaviors, and outcomes, and that personality inventories designed to capture pathological variance can be conceptualized as maladaptive extensions of the established normative factors. However, this relationship is untested in the conditioning literature. At present, it

is unknown if normative, pathological, or combined measures will best predict fear and avoidance generalization or APIC processes.

For a clinical example of why these gaps are in need of address, consider the following: there is evidence that people with PTSD overgeneralize compared to those without PTSD (Kaczurkin et al., 2016; R. A. Morey et al., 2015), yet we still know extremely little about generalization differences between people who have PTSD (i.e., within-group variance) because the prior studies primarily utilized a group-averaging approach (i.e., between-group variance/ mean differences). We also know that PTSD is a highly heterogeneous disorder in terms of symptom profile (Contractor, Roley-Roberts, Lagdon, & Armour, 2017; DiMauro et al., 2014; Zoellner, Pruitt, Farach, & Jun, 2014), personality traits (Contractor, Armour, Shea, Mota, & Pietrzak, 2016; Thomas et al., 2014), and even neurobiological responsivity (Lanius, Bluhm, Lanius, & Pain, 2006). Therefore, we have somewhat of a “black box” situation due to the means used to characterize the groups and test between-subject variance. Broadly speaking, this is not an issue if 1) we assume everyone in the PTSD group is extremely similar on all variables of interest and we have accounted for all vital covariates for the aims of the current study (which, even in the most well-controlled study, is a difficult goal to achieve) and 2) that generalization did not meaningfully differ between those in the PTSD group. This becomes problematic when we consider 1) other variables that are very relevant to conceptualization, diagnosis, and treatment, but were not part of the previous studies that we know are sources of heterogeneity in PTSD (e.g., impulsivity, anhedonia), and 2) that we know as a point of fact that generalization did differ between people in the PTSD group in the cited studies, but it is not clear what proportion of this variance is due to

measurement or random error and what is due to personality variation that might covary with generalization indices. Further, we as of yet have no empirical evidence of instrumental avoidance in PTSD, and thus have not observed how fear generalization influences generalized avoidance. To tie this back into the clinical picture, we know nothing about what predicts differences in fear generalization amongst people with PTSD, a primary target for PTSD treatment (for reviews, see S.A. Rauch, Eftekhari, & Ruzek, 2012; Sripada, Rauch, & Liberzon, 2016), nor do we know how generalized fear predicts avoidance in people with PTSD (with generalized avoidance being another primary target of PTSD interventions), and finally we also know that people with the PTSD diagnostic label greatly vary on a large number of clinically relevant personality traits. In short, we know very little that is helpful for the clinician who is treating an individual that, besides being diagnosed with PTSD, might deviate considerably from the “average” participant in the lab studies. Yes, the clinician is further convinced by the empirical literature that overgeneralization is contributing to his clients symptoms and is confident in their choice of a treatment that addresses this and other pathological processes, but the empirical literature is silent in regards to the myriad of determinants that contribute to this particular client’s problems and how they can be addressed in the clinic. The previously listed gaps in the research are all, to some extent, related to issues brought up in this clinical example, and of course these issues do not just relate to clinical examples: as the literature currently stands, we would not be able to confidently predict if the child who was scared of the Rottweiler will develop pathology, and if he does we would not have a good idea of what particular traits contributed to that development. Clearly, there is much needed work in this area that has yet to be done.

The Current Study and Specific Aims

In the current study, we explicitly address the outlined gaps in the literature by applying a broad individual differences approach to the study of fear generalization, avoidance generalization, and the covariation between the two using a previously validated paradigm (van Meurs et al., 2014) to test a large sample of non-patient participants. This task yields indices of both Pavlovian fear and instrumental avoidance covariation (APIC) in the context of maladaptive avoidance of benign generalization stimuli (APIC-G) and more adaptive avoidance of conditioned threat cues (APIC-CS+), making it ideal for addressing the stated gaps. To ensure we comprehensively assess sources of variation that could explain individual differences in these dependent variables and are relevant as potential risk or protective factors for psychopathology, we use measures covering a large range of human variance (e.g., Big Five personality traits; e.g., Costa & McCrae, 1992; Goldberg, 1990) to assess broadband personality traits that are both conceptually related to fear and avoidance conditioning (e.g., Neuroticism). We also assess broadband traits that are not necessarily closely, directly, or intuitively linked to fear and avoidance but might contribute to or modify the expression of fear and avoidance (e.g., Extraversion and Conscientiousness), as well as their generalized forms, within the context of an approach-avoidance conflict. These traits are assessed in conjunction with narrowband measures that are more narrowly related to symptoms of fear and anxiety pathology (e.g., trait anxiety and fear, anxiety and sensitivity intolerance of uncertainty). Further, we include measures of both normative (e.g., Big Five personality traits) and pathological (e.g., DSM-5 maladaptive personality traits) broadband traits, lower-level factors (e.g., aspects) of these traits, and use empirically-

supported composite indices based on both the normative and pathological measures to ensure we are 1) testing a broader range of variance to avoid potential limitations from prior work and 2) including the extremes of personality variation that were likely driving results in studies that successfully found generalization effects in patients with anxiety and trauma disorders. Finally, we designed our analytic plan around a multilevel modeling framework to facilitate a dimensional individual differences approach that yields improved statistical fidelity and precision, which represents a needed step forward for the field. Overall, we hope this work can both provide new or improved insight into basic processes investigated by clinical translational scientists, as well as provide a methodological and statistical foundation that facilitates future basic science work.

Specific aims.

The overall aim of this study is broad and therefore requires more detailed operationalization. Thus, we have identified four primary specific aims for this study. We aim to 1) identify dimensional personality predictors of generalized Pavlovian conditioned fear, as measured both physiologically and behaviorally, and generalized instrumental avoidance, as measured with overt behavior; 2) test whether these personality predictors significantly and meaningfully improve prediction of instrumental avoidance in APIC models; 3) identify which personality predictors moderate APIC either through facilitation or attenuation (i.e., dispositional risk and protective factors) of the relationship; and 4) contextualize results in terms of relative adaptiveness of the approach-avoidance conflict outcome, which is accomplished through follow-up testing of APIC-G (maladaptive avoidance) and APIC-CS+ (adaptive avoidance) using reduced versions of the APIC model (i.e., with fewer stimulus classes included in the model). In

the following section, we expand on these four aims in the context of specific predictions. Note that to facilitate readability, generalized Pavlovian conditioned fear will henceforth in this dissertation be referred to as fear generalization or generalized fear and generalized instrumental avoidance will be referred to as generalized avoidance or generalized avoidance.

Hypotheses and exploratory testing approach.

Due to a wealth of research findings with direct relevance to the current investigation, numerous *a priori* predictions are justified. What follows is a relatively in-depth description of each aim and the associated testing strategy, then specific hypotheses (when possible) for that aim which are ordered sequentially by each dispositional trait (i.e., we discuss Neuroticism predictions first, then Conscientiousness, etc.). However, in certain cases, 1) there is not sufficient published research available; 2) there are notably contradictory or inconclusive findings which are exacerbated by methodological heterogeneity; or 3) an argument based on solid theoretical rationale is not available. In these cases, specific *a priori* hypotheses are not tenable, and an exploratory approach is required and represents an appropriate step forward in the empirical tradition (e.g., Jebb, Parrigon, & Woo, 2017). If personality variables without a corresponding hypothesis are found to be significant predictors and/or moderators, we will report and interpret those findings and note their exploratory nature.

Aim 1 (A1) Identifying personality predictors of generalization.

This aim encapsulates one of the primary contributions of this dissertation: identifying which personality traits (both narrowband and broadband, normative and pathological) are associated with performance variables on a task measuring fear and

avoidance generalization. Accomplishing this aim includes testing and interpretation of main effects (i.e., does a trait significantly predict the magnitude of an outcome variable across all trial-types) and, most importantly, interaction effects. There are two types of interaction effects being tested: linear and quadratic interactions. Both types of interactions provide insight into potential associations between personality traits and generalization processes; therefore each significant interaction will be interpreted and graphed with the goal of identifying generalization effects. Overall, this aim represents a first comprehensive step in applying personality methods to the study of fear and anxiety generalization. All narrowband traits are tested separately for this aim and all other analyses. See Appendix A for specific details regarding the series of models used to test broadband traits this aim.

The majority of our specific hypotheses for this aim focus on measures of negative affect (e.g., Neuroticism, its aspects, and related personality constructs; narrowband anxiety- and fear-specific variables), as the majority of the literature did not include other types of traits as covariates in their predictive models. Therefore, for this aim we began all lines of analyses by constructing separate models in which Neuroticism or the equivalent variable from another scale (Negative Affectivity and “Distress-PB”, or the narrowband variables) or its aspects (Withdrawal, Volatility) was the only trait predictor (both in main effect and interaction models), which allows us to more clearly compare the current study’s results with prior individual differences in fear generalization work. However, the scope of this study, and this aim in particular, is broader than negative affect variables, and we thus constructed models for the broadband traits that include both the relevant negative affect trait and the other traits of the same kind and

level (e.g., Neuroticism with Conscientiousness and Extraversion, Negative Affectivity with Disinhibition and Detachment, Volatility and Withdrawal with the other four aspects tested in this study, etc.), which enable us to test the unique predictive contributions of the negative affect variables, as well as the other personality variables included in the current study.

Given that the current literature cannot strongly support predictions regarding other personality variables (e.g., Extraversion, Conscientiousness), we refrain from testing these variables in their own models (i.e., models that only contain one personality variable of interest). An exception is for models predicting avoidance: due to the empirical evidence and theoretical writings regarding how Conscientiousness and Extraversion (as well as their aspects and their corresponding pathological manifestations, Disinhibition and Detachment) operate more broadly on behavior, we initially included non-negative affect variables in their own models when predicting avoidance, and provide specific hypotheses for these models.

Specific hypotheses to address this aim, along with rationale for each hypothesis, are as follows:

- A1.H1: Withdrawal, an aspect of Neuroticism, will be associated with greater generalization across measures, whereas Neuroticism itself and Volatility, the other aspect of Neuroticism, will not have a significant relationship with generalization. We hypothesize this based on Withdrawal representing most of the “anxiety” variance in Neuroticism, as opposed to the “anger/irritability” variance that is represented by Volatility (DeYoung et al., 2007) and therefore Withdrawal is the most likely of both aspects and the overall trait to predict

generalization. That said, Withdrawal also contains items related to depression, which might attenuate the hypothesized relationship. Finally, we predict that these associations will persist when controlling for other personality variables.

- A1.H2: Negative Affectivity and the Neuroticism + Negative Affectivity composite, (Distress-PB), will also be associated with greater generalization across measures. Both of these variables are likely better approximations of the disorders used in studies identifying shallow generalization gradients in anxiety and trauma pathology than the normative traits described above. As with the previous hypothesis, we predict that these associations will persist when controlling for other personality variables.
- A1.H3: Conscientiousness will be associated with decreased generalized avoidance. This follows from its status as a “protective” trait that assists in goal-directed behavior and precision (Roberts et al., 2014a), which might protect against poorer discrimination between threat and safety signal value that underlies generalization processes. However, extremely high Conscientiousness is associated with pathological outcomes, including OCD, and we therefore hypothesize that Orderliness will best capture this pathological extreme and be associated with increased generalization. Further, we expect Industriousness will continue to account for the positive benefits of Conscientiousness and be associated with less avoidance generalization. We refrain from specific hypotheses regarding overall vs unique contributions of these variables (i.e., when they are modeled in isolation vs. with the other

personality traits/aspects), with the exception that Orderliness will be a stronger predictor of generalized avoidance when controlling for Industriousness.

- A1.H4: As with the previous hypothesis, we predict Extraversion will be associated with decreased generalized avoidance, as increased Extraversion will correspond with greater reward sensitivity and approach tendencies that will buffer against avoidance tendencies. Further, we predict this effect will be driven by Assertiveness, which contains content that seems the most relevant to approach under conditions of risk. We refrain from specific hypotheses regarding the overall vs unique contributions of these variables.
- A1.H5: In terms of narrowband traits, we predict that higher levels of TF-44, IUSF, and ASI will all be associated with increased generalization. We predict a null result for the STAI-T, given inconsistent past results and concerns about its construct validity. Instead, we believe the three aforementioned narrowband variables will be better high-resolution predictors of generalization than STAI-T.

Aim 2 (A2) Testing improved prediction of avoidance in APIC models.

This aim centers around the concept of improving the parametrization of our statistical models in the service of improved prediction of avoidance – if we know a person’s level of fear in response to the experimental manipulation, what other pieces of information will help us predict if the person will avoid? For example, if we have reasonably accounted for sources of fear and anxiety (via Pavlovian experimental indices and self-reported indices of fear, such as trait fear), then does knowing a person’s level of

Conscientiousness or Extraversion sharpen our prediction? This aim will broadly address this question and uses the same sets of *a priori* defined individual difference predictors that were used within the avoidance models in Aim 1 (see Appendix A). Additionally, separate models are constructed for testing APIC with Pavlovian responding operationalized by startle and by risk ratings, resulting in a parallel series of models. For this aim, we are interested in comparing models to determine if there is incremental improvement in models with personality traits of interest (i.e., does the APIC model that includes the personality trait provide significantly better prediction than the base APIC model?). Therefore, the following specific hypotheses do not refer to individual effects (main effects or interactions) and instead are structured around differences in predictive models. Specific hypotheses to address this aim, along with rationale for each hypothesis, are as follows:

- A2.H1: We predict that all of the higher-level personality traits being tested (Neuroticism, Negative Affectivity, and the composite using components from both, Distress-PB) will significantly improve prediction of avoidance when added to an APIC model. We base the overall prediction on the observation that Pavlovian variables significantly, but not strongly, predict avoidance and therefore there is likely sufficient variance related to fear “left over” for the trait variables to provide added predictive power. In addition, our prediction differs slightly from Aim 1, in which we did not predict Neuroticism would relate to generalization – we predict that Neuroticism will be suppressed by the Pavlovian variable and therefore become a stronger predictor of avoidance. However, as opposed to Aim 1, we do not make specific

predictions regarding the Withdrawal aspect given that Withdrawal contains depression variance, which possibly differentially relates to both the predictor (Pavlovian variable) and outcome (avoidance) in this model and makes it difficult to provide precise predictions.

- A2.H2: We predict Conscientiousness will also significantly improve prediction of avoidance when added to an APIC model, as it provides a source of unique variance related to the task (tendencies towards goal obtainment and rule-following) that is not captured by other parts of the base APIC model. We also predict that modeling Industriousness and Orderliness separately in the same model will result in a better model fit than a model with just Conscientiousness; this follows from the documented suppression effect between Industriousness and Orderliness.
- A2.H3: We predict Extraversion will also significantly improve prediction of avoidance when added to an APIC model, as it also provides a source of unique variance related to the task (reward motivation and approach tendency) that is not captured by other parts of the base APIC model.
- A2.H4: In terms of narrowband traits, similar to Aim 1 we predict that TF-44, IUSF, and ASI will all significantly improve prediction when added to (separate) APIC models. However, in contrast to Aim 1, we predict a significant effect of STAI-T, such that it also will contribute significant predictive power to the base APIC model. We make this prediction for the same reason we predicted Neuroticism would significantly improve

prediction: a suppression effect will occur between the Pavlovian variable and STAI-T.

Aim 3 (A3) Moderation of APIC by dispositional variables.

This aim is concerned with identifying specific dispositions that moderate the relationship between Pavlovian variables and instrumental avoidance across the complete stimulus continuum (i.e., APIC), either by facilitating the relationship (i.e., positively moderate or strengthen the association between the Pavlovian variable and avoidance) or attenuating the relationship (i.e., negatively moderate or weaken the association between the Pavlovian variable and avoidance). As for Aim 2, separate models are constructed for testing APIC with Pavlovian responding operationalized by startle and by risk ratings, resulting in a parallel series of models. Unlike Aim 2, this aim focuses on individual effects of personality variables, specifically by quantifying the degree to which these variables moderate APIC and the pattern of moderation effects across different stimulus levels. Statistically, this involves testing the predictive properties of a three-way interaction involving the personality variable of interest, a Pavlovian variable, and the Stimulus dimension. As the goal of this aim is to broadly identify APIC moderators, the entire stimulus continuum (both CS-s, all GSs, and the CS+) are included in these models. Hypotheses to address this aim, along with rationale for each hypothesis, are as follows:

- A3.H1: We predict that the negative affect personality variables with pathological variance (Negative Affectivity and Distress-PB) will significantly moderate APIC. We based this prediction on the theory and limited human evidence that APIC is primarily a pathological process

which “converts” fear and anxiety into maladaptive avoidance, which makes up the majority of APIC variance. This is also the reason we do not make a specific hypothesis for Neuroticism, and contend that Neuroticism and its aspects are less likely to moderate APIC.

- A3.H2: We predict Conscientiousness will be a significant moderator of APIC, but a weak one unless decomposed into its component aspects. This follows from the opposing motivations represented by Industriousness and Orderliness and how they likely relate to APIC. We predict that higher levels of Orderliness will facilitate APIC relations, as those who are higher on this aspect might be more inclined towards controlling internal sensations in line with the associated behavioral drive (e.g., organizing a room when bothered by its state of messiness to reduce distress) and therefore more likely to avoid as fear increases. Conversely, we predict that higher levels of Industriousness will weaken APIC relations, as those higher on this aspect might be more likely to persevere in the service of obtaining a goal, and therefore less likely to avoid due to the consequence of forgoing a “win”.
- A3.H3: We predict Extraversion will be a significant moderator of APIC, with lower Extraversion facilitating APIC. This follows from Extraversion generally corresponding to approach motivation, and that those with lower approach motivation will be less inclined to persevere (i.e., approach) while afraid in service of obtaining a reward. We also predict that the Assertiveness aspect will largely drive this moderation effect, as despite

the socially-focused item content, this lower-level factor has been related to general personal agency and goal-orientation (e.g., Depue & Collins, 1999) and potentially represents a protective factor, such that those lower on Assertiveness will demonstrate greater APIC.

- A3.H4: Consistent with previous aims, we hypothesize that of the narrowband variables, TF-44, IUSF, and ASI will facilitate APIC, but not STAI-T. In line with their conceptualizations as indicators of internalizing pathology, we predict that higher levels of these three traits will be associated with increased APIC.

Aim 4 (A4) Specific APIC-G and APIC-CS+ effects.

Aim 4 is nearly identical in goals and methodological approach to Aim 3; however, testing is limited to APIC-G and APIC-CS+ models so as to focus on maladaptive or adaptive avoidance without controlling for the other. This is done because 1) the full APIC model might obscure relevant personality moderators with modest effects on either APIC-G or APIC-CS+, but not both; 2) to determine if identified APIC moderators persist when tested within the APIC-G and/or APIC-CS+ models; and 3) to approximate the analytic approach from previous conditioning work with the goal of facilitating interpretation of this dissertation's results in the contexts of that prior work. As previously noted, the APIC-G/APIC-CS+ literature in humans is practically non-existent, and there are no available studies of personality moderators. Therefore, we provide a limited set of broad predictions, and contend that it is necessary to treat this aim as a primarily exploratory endeavor.

- A4.H1: Similar to Aim 3, we predict that the negative affect personality variables with pathological variance (Negative Affectivity and Distress-PB) will significantly moderate APIC-G, but not APIC-CS+. Again, our rationale is predicated on the evidence that generalization is strongest in those with pathological anxiety conditions, that a primary factor underlying the pathology is APIC-G, and that these personality traits will be the most appropriate analogues for these conditions in the current study.
- A4.H2: We hypothesize APIC-CS+ will be moderated by Conscientiousness, consistent with the equivalent hypothesis for Aim 3. Specifically, we expect higher levels of Industriousness to buffer against conversion of fear responding into avoidance for the danger cue, and the opposite for those higher on Orderliness. We limit this prediction to APIC-CS+ due to this being the least ambiguous approach-avoidance conflict contained within current experimental paradigms testing APIC (i.e., provokes the clearest conflict between obtaining a goal, reducing distress/harm, and “following the rules”) and therefore that the situational demands of CS+ trials are the most likely to activate Conscientiousness-related individual differences.
- A4.H3: We hypothesize that lower Extraversion (driven by the Assertiveness aspect) will significantly facilitate APIC-G, and again consistent with Aim 3, contend that this will be due to lower approach motivation that will decrease the chance of taking the risk of approaching during an ambiguous situation (GS trials), even when the stimuli only elicit a low or moderate degree of fear. We do not predict this moderation effect will hold for APIC-CS+, as the higher degree of

fear experienced during the CS+ will not necessarily be offset by the normative degree of approach motivation seen in the Extraversion trait and its aspects. Also of note is the possibility that these traits might be capturing the same variance (i.e., the same participants) who could drive a significant Negative Affect or Distress-PB moderating effect for APIC-G – those generally high on negative affect traits are also those likely lower on positive affect trait (L. A. Clark & Watson, 1991).

Method

Participants

We recruited and tested 396 undergraduate and graduate students from the University of Minnesota-Twin Cities campus and the surrounding community. All participants were English-speaking adults between ages 18 and 50 who are enrolled in an undergraduate psychology course and receive research credit for study participation. Criteria for exclusion from initial testing were: 1) vision or hearing conditions that could interfere with completion of the experimental task; 2) active use of antipsychotics, benzodiazepines, or tranquilizers; 3) use of alcohol within twenty-four hours of study start; 4) or use of nicotine or caffeine within 3 hours of study start; or 6) active suicidal or homicidal intent. Due to the aims of the study involving recruitment of a broad and representative sample, current or past psychopathology was *not* an explicit exclusion criterion, although exclusion based on certain medications (e.g., antipsychotics) might have *de facto* excluded participants with certain forms of psychopathology.

Of the 396 participants tested for the study, a total of 108 were excluded from final analysis, leaving a final sample size of 288 (65.97% female, $M_{\text{age}} = 20.47$, $SD_{\text{age}} = 3.32$). See Table 1 for a demographic overview for the current sample. Despite losing ~28% of our data, the final sample subjected to analysis is among the largest samples tested in a conditioning study (Lonsdorf & Merz, 2017). The most common reasons for exclusion were that the participant did not complete the task due to technical error or they elected to discontinue, they did not demonstrate discriminative fear conditioning (e.g., responses to the CS+ were equal to or lower than CS- responses during the Pavlovian trials of generalization and they did not “learn the task”), questionable self-report questionnaire validity based on embedded items created for this purpose, or there were technical difficulties that resulted in an unusable dataset (e.g., sufficient physiological data was not collected). See Table 2 for exclusion conditions and counts for each category and specific conditions.

Table 1. Participant demographics (N = 288)

<u>Variable</u>	<u>N</u>	<u>%</u>
Female	190	65.97%
Ethnicity		
White (non-Hispanic)	190	65.97%
Asian	66	22.92%
Other or Multiple/Mixed-race	13	4.51%
African American or Black	9	3.13%
Hispanic	7	2.43%
Middle Eastern	2	0.69%
Unknown/Did not answer	1	0.35%
	Mean	SD
Age (years)	20.47	3.32
Education (years)	15.04	1.9

Note: Participants self-identified gender and ethnicity, with open-ended input available for both questions to allow for responses that

did not correspond with provided options. All participants in the reported sample self-identified as either female/woman or male/man.

Table 2. Exclusion conditions and total excluded

Exclusion condition	N
Performance	35
Did not learn CS+/CS- contingency	29
Avoidance outlier	3
EMG non-responder ($<1 \mu\text{V}$)	3
Technical issues	51
Missing/unreliable EMG	33
Missing/unreliable risk ratings	13
Other technical issues	5
Other	45
Validity concerns on questionnaires	26
Current antipsychotic use	1
Explicit effort concerns during testing	1
Withdrew during lab visit	17
Total, ≥ 1 exclusion condition met	131
Total, ≥ 2 exclusion conditions met	23
Total, participants excluded	108

Note: Statistical outliers for mean avoidance rate were all at 100% avoidance. Evaluation of CS+/CS- contingency based on risk ratings during the generalization phase. CS- = conditioned safety cue; CS+ conditioned danger cue; EMG = electromyography; μV = microvolt.

Materials

Trait-level questionnaires

All trait-level questionnaires were administered via an online system that was self-paced and completed prior to a laboratory visit completed soon after these questionnaires. We embedded a total of three “validity” items within longer

questionnaires (no more than 1 in a questionnaire) to assist in the detection of inattentive and/or idiosyncratic responding. Example of validity items include “The University of Minnesota is located in Wisconsin” and “I was born in 1876”.

Broadband personality measures.

Big Five Aspect Scale (BFAS; DeYoung, Quilty, & Peterson, 2007).

The BFAS is a factor-analysis derived measure of the Big Five traits (Openness, Conscientiousness, Extroversion, Agreeableness, Neuroticism) and the immediately subordinate aspects of each trait (2 for each domain, for a total of 10 domains). The BFAS is a more recently developed personality inventory that built on research done on existing and commonly used Big Five measures (Costa & McCrae, 1992; Goldberg, 1999) and, through its focus on aspects, captures a level of the Big Five structure that exists between the broad traits and the more specific and narrow facets. DeYoung et al., (2007) also propose that the BFAS factor solution provides a more optimal classification structure for biological investigations of personality (i.e., the aspects can be characterized by nominally separate biological substrates and correspond to genotypic variation; Jang, Livesley, Angleitner, Riemann, & Vernon, 2002). Accordingly, the BFAS is an ideal personality measure for the current investigation due the use of neurobiologically informed experimental paradigms (Pavlovian conditioning) and measures (e.g., fear-potentiated startle). In the current study, we calculated the trait and aspect scores in line with scoring method from DeYoung et al. (2007), in which each aspect is comprised of the mean of 10 aspect-specific items, and the traits are calculated as the means of its two component aspects. Our online system did not permit missing/skipped items on the BFAS

given the importance of personality data to the current study, and therefore there was no missing data in any calculated scales for participants in the final sample.

Personality Inventory for DSM-5 (PID-5; Krueger, Derringer, Markon, Watson, & Skodol, 2012).

The PID-5, as with the BFAS, measures broader personality domains, but is designed to capture maladaptive personality traits and all of the scales are thus keyed in the direction of higher scores corresponding to higher levels of psychopathology. The PID-5 indexes five higher order domains (Psychoticism, Disinhibition, Detachment, Antagonism, Negative Affectivity) and 25 lower-level facets. These facets load on one or more PID-5 trait, with some facets more clearly reflecting a single trait (e.g., the “Manipulativeness” facet as part of the “Antagonism” trait) than others (the “Depressivity” facet loads almost equally as strongly on the “Negative Affectivity” and “Disinhibition” traits as it does on the “Detachment” trait) (e.g., De Fruyt et al., 2013; Krueger, Derringer, Markon, Watson, & Skodol, 2012; Wright et al., 2012). The PID-5 is also conceptualized as a measure with explicit clinical utility, as it was developed as an accompanying measure for the DSM-5 alternative “hybrid” model of personality disorder (American Psychiatric Association, 2013; Krueger et al., 2012) and the PID-5 scales show good convergence with the conceptually related scales of personality inventories commonly used in clinical assessment settings (Anderson et al., 2013; Hopwood, Wright, et al., 2013), as well as scales specifically designed for personality disorder detection (Bastiaens et al., 2016; Few et al., 2013; Fossati, Krueger, Markon, Borroni, & Maffei, 2013; Hopwood, Thomas, Markon, Wright, & Krueger, 2012).

Given the five-factor structure of the PID-5, it has sometimes been referred to as a “maladaptive Big Five” (e.g., DeYoung, Carey, Krueger, & Ross, 2016; Gore & Widiger, 2013; Suzuki, Samuel, Pahlen, & Krueger, 2015; Thomas et al., 2013), which is a mostly supported statement; the exception is that the Psychoticism trait does not as precisely overlap with its conceptual pairing, Openness, as well as the other four trait pairs. That said, factor-analytic methods consistently find overall strong convergence between Big Five and PID-5 constructs in both nonclinical (Suzuki et al., 2015; Thomas et al., 2013) and clinical samples (Quilty, Ayearst, Chmielewski, Pollock, & Bagby, 2013; A. G. C. Wright & Simms, 2014). In summary, the PID-5 appears to capture the maladaptive “extremes” of the Big Five and that these measures are quantifying variance on the same set of dimensions.

In the current study, we calculated the trait and facet scores in line with the method used in the original development of the PID-5 (Krueger, Derringer, Markon, Watson, & Skodol, 2012) in which each facet comprised of the mean of a variable number of facet-specific items, and the traits are calculated as the means of the facets that load most strongly on them in prior work, (i.e., facets are allowed to contribute to only one trait/ traits do not share facets). Note that in this respect our method differed from Krueger and colleagues’ (2016) and the recommendations for the PID-5 measure that is published by the APA and reflects the method used in other studies using the PID-5 (e.g., Suzuki, Samuel, Pahlen, & Krueger, 2015). As with the BFAS, our online system did not permit missing/skipped items on the PID-5, and therefore there was no missing data in any calculated scales for participants in the final sample.

PID-5 and BFAS composite scales (“PB” scales).

Given the conceptual and empirical overlap between the PID-5 and Big Five personality structure, the two personality inventories used in the current investigation are well-suited to testing the full range of personality variance. Specifically, the BFAS yields a normally distributed, broad range of personality variance (DeYoung et al., 2007) that potentially includes some pathological manifestations at the tails, whereas the PID-5, which is not normally distributed in the general population (Krueger et al., 2012), overlaps most strongly with the BFAS tails at the lower end of its distribution, as it is designed to capture only maladaptive or pathological personality variance. Further, factor analytic studies of the combined PID-5 and BFAS by the respective authors of each inventory find that PID-5 facets and BFAS domains demonstrate a good fit with a 10-factor solution, with one BFAS domain loading highly on one factor (with each BFAS domain being the highest loading domain on only one factor) along with conceptually-linked PID-5 facets (e.g., DeYoung, Carey, Krueger, & Ross, 2016); the 10 factors generally resemble the 10 BFAS aspects. This provides support for combining the BFAS and PID-5 to capture a more expansive range of maladaptive and adaptive personality variance through construction of scales that match the factor analysis results from DeYoung et al. (2016).

To create these PID-5+BFAS scales (termed “PB” scales in this investigation), we used the factor structure derived from Sample 2 (which had both a larger sample than Sample 1 and was the only sample to exclude participants who did not respond correctly to “attention checks” similar to those used in the current study) in DeYoung et al. (2016) to identify the scales comprising the PB scales of interest, standardized all scales, and then averaged those standardized scales together to create the PB scales. Additionally, we

reversed-keyed PID-5 scales that negatively correlated with the PB scale (e.g., the “Submissiveness” facet on the PID-5 negatively correlates with the “Assertiveness” PB factor) to ensure the average score could be correctly interpreted.

Narrowband personality measures.

Trait Fear – 44 Item Questionnaire (TF-44; Kramer et al., 2019). The TF-44 is an empirically-derived dimensional measure of trait fear (when defining the overarching construct of “threat sensitivity” from the “fear” end of this bipolar dimension) that was originally developed through exploratory factor analysis of a large twin sample that was characterized with multiple internalizing and externalizing self-report measures and corresponding psychophysiological measurements (*Kramer, Patrick, Krueger, & Gasperi, 2012*). Importantly, the trait fear dimension as assessed by the TF-44 is positively associated with biological indices of defensive responding (e.g., fear-potentiated startle; *Kramer, Patrick, Krueger, & Gasperi, 2012*) and other measures of fear and threat sensitivity (e.g., *Kramer et al., 2019*). The TF-44 was scored in line with the method used in *Kramer et al. (2019)*, with a higher average on all items indicating a higher level of trait fear. We calculated a prorated score for cases in which ≤ 3 items were missing.

Intolerance of Uncertainty – Short Form (IUSF; Carleton, Norton, & Asmundson, 2007).

The IUSF is a dimensional, commonly used, brief measure of IU that converges well with the original measure. A two-factor solution fits the IUSF well and indicates the presence of distinguishable Prospective and Inhibitory types of IU in the measure. The IUSF and its subscales were scored in line with the method used in *Carleton et al. (2007)*

with a higher average on all items indicating a higher intolerance of uncertainty. We calculated a prorated score for cases in which ≤ 3 items were missing.

Anxiety Sensitivity Index (ASI; Reiss, Peterson, Gursky, & McNally, 1986).

The ASI is a dimensional, commonly used measure of AS. Later studies of the ASI proposed three subscales (Mental, Physical, and Social) based on a best-fitting hierarchical three-factor structure (Zinbarg, Barlow, & Brown, 1997). The ASI and its subscales was scored in line with the method used in Reiss et al. (1986) and Zinbarg et al (1997), with a higher average on all items indicating higher anxiety sensitivity. We calculated a prorated score for cases in which ≤ 3 items were missing.

Spielberger Trait-State Anxiety Inventory – Trait (STAI-T; Spielberger et al., 1983).

The STAI-T is a dimensional, commonly used measure of trait anxiety. As discussed earlier, the trait anxiety construct has received considerable criticism (e.g., Endler & Kocovski, 2001), and current efforts highlight the need to distinguish between trait anxiety and trait fear (Kramer et al., 2019; Sylvers, Lilienfeld, & LaPrairie, 2011). Nevertheless, it is included for across-study comparison purposes, as a large number of conditioning studies using the STAI-T/trait anxiety as their primary dimensional negative affect measure (Lonsdorf & Merz, 2017). The STAI-T was scored in line with the method used in Spielberger et al. (1983) with a higher average on all items indicating higher trait anxiety. We calculated a prorated score for cases in which ≤ 3 items were missing.

State level questionnaires

In addition to broadband and narrowband personality assessment, the current study includes brief measures designed to capture current symptoms associated with

anxiety, stress, and depression before the completion of the experimental procedures. Relevant measures include state anxiety (part of the STAI [STAI-S]; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983), perceived stress level (Perceived Stress Scale [PSS]; Cohen, Kamarck, & Mermelstein, 1994), and depressive symptoms (Beck Depression Inventory [BDI-II]; Beck, Steer, & Brown, 1996).

Physiological apparatus

We controlled stimulation and physiological recording via a commercial system (Contact Precision Instruments, London) and measured startle blink with electromyography (EMG) using two 6-mm tincup electrodes (sampling rate = 1000 Hz; online bandwidth filter = 30–500 Hz) applied to the *orbicularis oculi* facial muscle. In accordance with standardized guidelines for human startle studies (Blumenthal et al., 2005) we placed one electrode below the right lower eyelid in line with the pupil while in forward gaze and placed the second electrode approximately 2 cm lateral to the first. Additionally, we placed a 9-mm disk electrode on the anterior forearm to serve as a ground. We probed the startle blink with a burst of white noise (40 ms, 102 dB) with a near instantaneous rise time, that was presented binaurally through headphones.

Pavlovian-Instrumental Generalization (PIG) task.

A previously validated task developed by our group was used to assess Pavlovian generalization, instrumental generalization, and APIC-G and APIC-CS+ (van Meurs et al., 2014). The context of the task is a farming video game in which the participant is represented by an abstract farmer avatar on a bicycle. The farmer's goal is to successfully plant and gather crops before a flock of pesky birds destroys the crops. Complicating matters is that the farmer rides a bicycle on his way to the crops, and there is a chance of

painfully falling off the bicycle (i.e., receiving a shock) under certain conditions. Stimuli are displayed in the center of the display and provide information about potential risk.

Stimuli.

Stimuli for the task include eight rings of gradually increasing size, with extremes serving as conditioned danger (CS+) and conditioned safety cues (oCS-). The six rings of intermediary size, averaged into 3 “bins”, serve as generalization stimuli (GSs, with the GS1 the stimulus most dissimilar to the CS+ and the GS3 the stimulus most similar to the CS+, with the GS2 in-between), and create a continuum-of-similarity between the CS+ and the CS-. See Figure 2 for a graphic displaying the stimulus continuum and additional information about the task. For 50% of participants the largest ring was the CS+, and for the other 50% the smallest ring was the CS+. Additionally, we used triangles as “non-circular” conditioned safety cues (Δ CS-) to assess the degree to which fear might generalize to all things circular, but not triangular, and provides a control condition for cue-based responses. Two sizes of triangles were used (roughly aligning in size with the largest and smallest circular stimuli) and responses were average together to form the Δ CS- “bin”. Finally, on some trials no cue or shape was presented, also without risk of shock (NS- trials), which provides an index of contextual fear and avoidance without a discrete cue. Stimuli were presented pseudo-randomly, and no more than two of the same stimuli were presented consecutively at any point in the task. The US was a brief electric shock (3-5 mA, 100-200ms) administered to the wrist that was calibrated for each participant to be uncomfortable but not painful (see Procedure for shock work-up details).

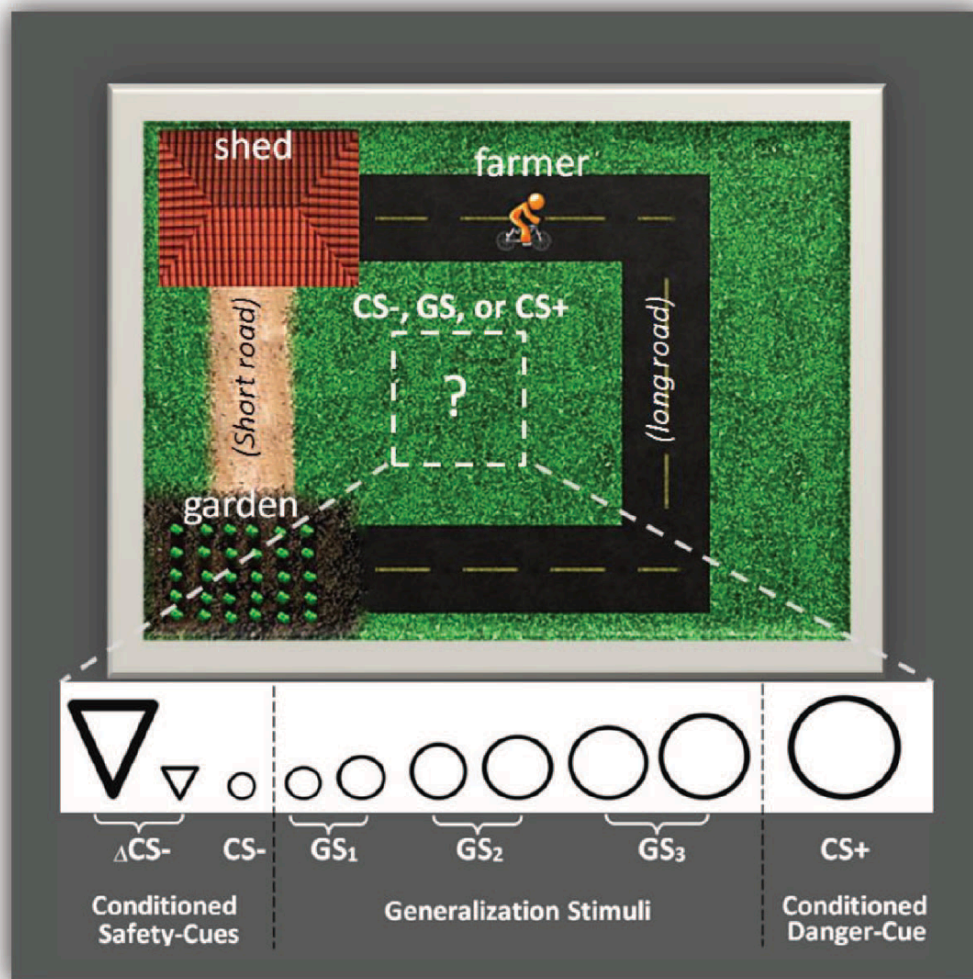


Figure 2. Sample graphic of the Pavlovian-Instrumental Generalization (PIG) paradigm. The upper part of the graphic shows the general context of the task (a farmer riding a bike to tend to his crops), the lower part shows the stimuli that appear in the center of the screen. For 50% of participants, the largest circle was conditioned as the danger cue and the smallest was the safety cue (as shown in the image); for the other 50% of participants this was reversed. Graphic adapted from a scientific presentation given by Dr. Lissek, with his permission. CS- = conditioned safety cue, GS = generalization stimuli, CS+ = conditioned danger cue.

Phases.

The PIG task is completed in three phases: *pre-acquisition*, *acquisition*, and *generalization*. In pre-acquisition and acquisition, the farmer can only travel across a short dirt path to reach the crops; a longer paved road is also present but is closed off for

construction and the participant does not have the option to take that path during these phases. The short dirt path is considered the “danger” path, as the farmer can only fall off his bicycle (receive a shock) on this path depending on which stimulus is presented. The longer paved road is the “safe” path, as there is no change of falling off the bicycle (receiving a shock) when on this path, regardless of the displayed stimulus. Only CS+, CS-, Δ CS-¹⁵, and NS- trials are encountered during these phases. In the pre-acquisition phase, shock is never administered/the farmer never falls off his bicycle and the participant is informed they are at no risk for shock. In the acquisition phase, participants are explicitly informed that they are at risk of shock and that if they attend to the task they can learn when shock will occur. During acquisition, shock is paired with the CS+, and is never administered on CS-, Δ CS-, and NS- trials. The contingency was reinforced on 100% of CS+ trials through a “shock” graphic representing the farmer falling off his bike. However, due to concerns about sensitization to the US and potential “strong situations” that reduce response variability (Lissek et al., 2006), the physical shock is administered on only 50% of CS+ trials. Due to the lack of ability to avoid the shock, these phases only contain trials that index Pavlovian conditioning.

During the generalization phase, there were two trial-types that alternated every-other trial: Pavlovian and instrumental trials. Pavlovian trials resembled those from the previous phases; although the longer paved road was now open, the participant did not have the option to take this road in Pavlovian trials and there remains a chance of shock if the CS+ is present (with the 50% physical shock reinforcement, 100% visual reinforcement maintained from the acquisition phase). These trials are framed as

¹⁵ Only the smaller sized triangle was presented during pre-acquisition and acquisition.

“planting” trials (the farmer is on his way to plant crops) and are procedurally the same as those in the acquisition phases except now GS trials can now be encountered, allowing for measurement of Pavlovian generalization.

On instrumental trials, the participant was alerted that they had 5 seconds to choose a path based on the displayed stimulus (or lack of stimulus). The participant was then able to choose a path via button press (1 = short dangerous path, 2 = long safe path). When taking the short path, the participant was always able to successfully gather the planted crops (win condition), but when taking the longer path there was a 75% likelihood that the participant would have the crops destroyed by pesky birds (lose condition). If participants did not decide within the allotted 5 seconds, they were forced to take the short path (and risk shock), yet now could not win by harvesting the crops. There was no other reward motivation for taking the short path and harvesting crops other than intrinsic motivation to do well on the task and a graphic of the farmer successfully harvesting the crops with a small shower of shimmering sparkles. Also important to note is that immediately prior to the generalization phase, participants completed a brief practice session to learn that the long path was now “open” and to receive instructions on how to select a path via button press. See Figure 3 for a graphical representation of the acquisition and generalization phases.

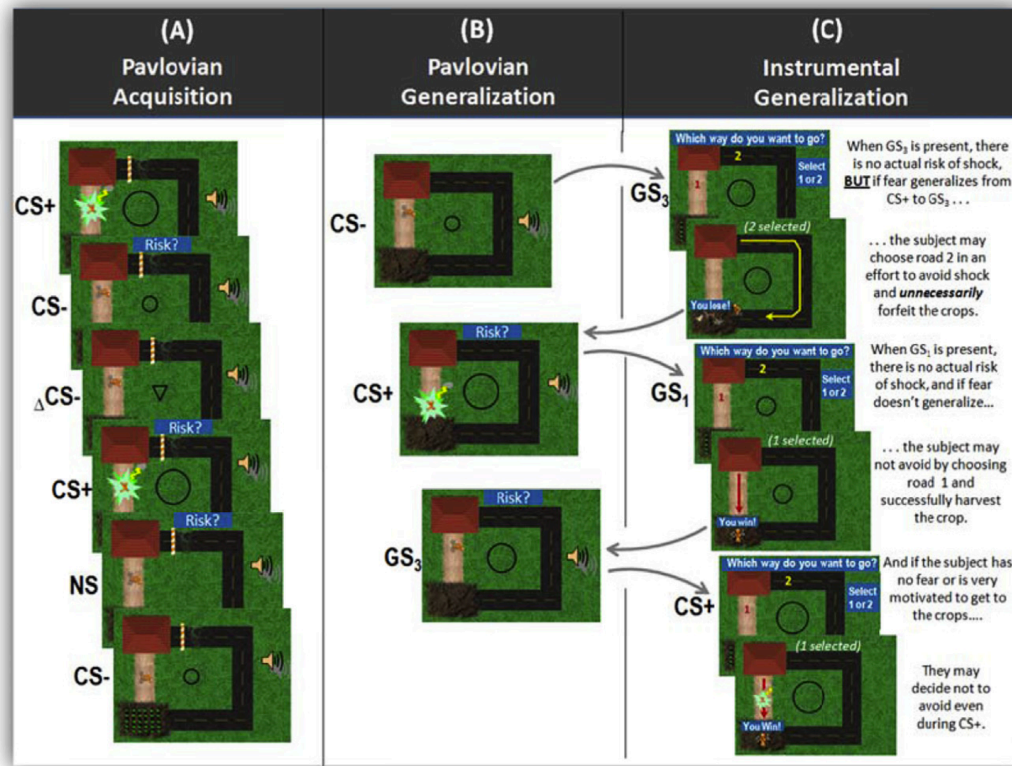


Figure 3. Example trials from the acquisition phase and generalization phase (both Pavlovian and Instrumental trials) of the Pavlovian-Instrumental Generalization (PIG) paradigm. Each graphic of the farm and roads is a distinct trial, except for the rightmost column, in which each pair represents a distinct trial. The leftmost column shows examples of a series of acquisition trials and how administration of shock is dependent on the presented stimulus. The center and rightmost column depict the progress of trials from Pavlovian, to Instrumental, to Pavlovian, and so on, and how participants can choose to approach or avoid. The blue “risk” box represents the prompt for participants to provide a risk rating on this particular trial. The small speaker graphic indicates a startle probe was administered on that trial. Graphic adapted from a scientific presentation given by Dr. Lissek, with his permission. CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimuli; NS = no shape.

Pavlovian response measures.

Pavlovian fear responses were measured both physiologically and behaviorally for each trial type. Using EMG and acoustic startle probes, we measured fear-potentiated startle, the reliable magnification of the startle reflex when an organism is in a state of fear (e.g., Davis, 1992). Startle was assessed while the farmer biked down the road on every Pavlovian trial during generalization (6 out of 6 trials per trial type) and was never

assessed during instrumental trials; this helped to slow habituation to the startle probes. This yielded a total of 36 startle measurements per participants (6 trials each for 6 trial types) during generalization. The startle probe was administered prior to shock, if shock was also administered on the trial (i.e., during CS+ trials). Greater fear-potentiated startle was used a physiological index of greater fear responsivity to a specific stimulus.

Perceived risk of shock (“risk ratings”) for a given trial was measured by an online button press when prompted (1 = no risk, 2 = some risk, 3 = high risk). We explicitly instructed participants to rate their perceived risk of shock in the given moment based on what was on the screen, and not risk for an acoustic startle probe. Risk ratings were collected on 3 out of 6 Pavlovian trials for each stimulus during the generalization phase and were never collected on instrumental trials. This yielded a total of 18 risk rating measurements per participants (3 trials each for 6 trial types) during generalization. Participants were instructed to provide their response as quickly as possible, and response time was also recorded, but not analyzed for the current dissertation and will be reported elsewhere. Higher risk ratings were used a behavioral index of higher perceived risk (i.e., threat estimation).

Instrumental response measure.

Instrumental responses were indexed by the dichotomous decision made on instrumental trials. When indexed at the individual trial level, responses were coded as 0 = approached and 1 = avoided. Trials in which the participant did not respond within the allotted 5 seconds were not coded and excluded from analyses. Instrumental response was recorded for all instrumental trials (6 out of 6 trials per trial type) and was never assessed during Pavlovian trials. This yielded a total of 36 instrumental decisions per

participants (6 trials each for 6 trial types). As with risk ratings, participants were instructed to provide their decision response as quickly as possible, and response time was recorded but not analyzed.

Procedure

Participants provide informed consent prior to completing the online battery of questionnaires. Soon after completing these questionnaires, but on a different day, participants arrived at our laboratory. Participants then complete measures of state variables (e.g., state anxiety, depressive symptoms) and brief cognitive testing (which is beyond the scope of this dissertation and will be reported elsewhere). Next, participants were informed that they will begin the part of the study that involves shock, and psychophysiological measurement and shock electrodes were attached using a standardized procedure. We then calibrated shock intensity to an appropriate level for each participant through a commonly used standardized “shock work-up” procedure. In this procedure, we asked participants to rate a series of sample shocks on a scale of 1–5 (1 = no discomfort/ pain, 5 = very painful), with the goal of finding the level of shock corresponding to a “3” to “3.5” rating (uncomfortable but not painful) for the participant. Once an appropriate level of shock was confirmed, participants put on headphones and were introduced to the PIG task.

The PIG task was presented on a desktop computer with a 22-inch monitor in a dimly lit experimental booth, with participants sitting approximately 26 inches away from the screen. Next, participants completed a brief habituation sequence in which the screen was blank and 9 startle probes were administered pseudo-randomly and presented with 9-22 seconds between each probe. After the habituation sequence, we provided

general instructions regarding the PIG task, and specifically informed participants that if they attended to the presented shapes they could learn to predict when a shock will occur, but were not explicitly told of the CS+/US contingency. Prior to the generalization phase, participants were instructed as to the safety/danger value of each path and are informed that what they learned from previous phases still applies. Participants complete questionnaires after the acquisition and generalization phase to assess conditioning and motivations for avoidance.

Data Processing and Statistical Analysis Plan

Startle data reduction and preparation.

To quantify fear-potentiated startle, we rectified and smooth startle EMG (20-ms moving window average). The onset latency window for the blink reflex was set at 20–100 ms, and the peak magnitude was determined within 120 ms of response onset. We then subtracted the average baseline EMG level for the 50 ms preceding the startle stimulus from the EMG peak levels, resulting in a final startle magnitude value represented in microvolts (μV). Zero magnitudes were included due to part of the startle reflex involving a degree of non-responsiveness; however, participants with mean startle $< 1 \mu\text{V}$ were excluded from analyses as non-responders.

Self-report data reduction and preparation.

Personality variables demonstrating skewness greater than .75 were log-transformed to normalize the distribution of said variables. Standardization of personality variables is described within individual analytic subsections that immediately follow this section, as the method of standardization differed based on analytic need.

We created five of the aforementioned “PB” scales, which were derived from the PID-5 and BFAS. The composite variables constructed were “Distress-PB” (contains Withdrawn Distress [BFAS]; Anxiousness, Depressivity, Separation Insecurity [PID-5]); “Industriousness-PB” (contains Industriousness [BFAS]; Distractibility, Impulsivity, Perseveration, Risk Taking, Irresponsibility [PID-5, reverse scaled]); “Orderliness-PB” (contains Orderliness [BFAS]; Rigid Perfectionism [PID-5]), “Enthusiasm-PB” (contains Enthusiasm [BFAS]; Social Withdrawal, Anhedonia, Restricted Affectivity, Intimacy Avoidance, Suspiciousness [PID-5, reverse scaled]); and “Assertiveness-PB” (contains Assertiveness [BFAS]; (Submissiveness [PID-5, reverse scaled])). All PB variables have this abbreviation appended to the end of the scale description to differentiate these scales from the BFAS scales that have the same name.

Preliminary statistical analyses.

Statistical analyses consisted of three phases: manipulation checks, sample characteristics, and main analyses. All analyses were conducted in R (R Core Team, 2014) in the R Studio environment (RStudio Team, 2018). Each analytic phase is described in detail below:

Manipulation checks.

We used rmANOVA models fitted in the ez package (Lawrence, 2013) to confirm that the PIG task was resulting in successful fear conditioning and generalized fear and avoidance that was consistent with prior results using this task. Each model contained a single Stimulus factor (Δ CS-, oCS-, CS+, and for the generalization phase, all GSs) modeled as a within-subjects factor, and the dependent variable was represented as the average response to each stimulus condition. Two separate models were run for both pre-

acquisition and acquisition experimental phases (one each for startle and risk rating data) and three were run for generalization (startle, risk rating, and avoidance models). In addition to testing of the within-subjects main effect, we conducted quadratic trend analysis to confirm generalization effects. To align with prior work from our group and others, within-subjects T-scores were used for analysis of startle data in rmANOVAs to control for potential between-subject variability resulting from non-psychological processes (e.g., obligatory startle) and sum of squares calculation was set to be identical to what was used in prior studies (i.e., Type 3 calculation – the SPSS default). We conducted Mauchly's sphericity test for all models and applied Greenhouse-Geisser correction to all tests that violated sphericity. Omega-squared (ω^2) effect size estimates are provided for all effects (Olejnik & Algina, 2003).

Sample characteristics.

We conducted two sets of zero-order correlation analyses to assist with characterizing our sample and preparing for interpretation of main analyses. The first set of correlational tests were those testing the associations amongst the measured dispositional individual difference traits. The second set of tests correlated each measure collected during generalization (startle, risk ratings, and avoidance) for each stimulus with each other; data from the pre-acquisition and acquisition phases were excluded from these analyses due to the foci of the current investigation.

Main statistical analyses.

All main analyses (i.e., those conducted to address the stated aims of this study) were conducted using multilevel models, which we fitted with the lme4 package (Bates, Mächler, Bolker, & Walker, 2015). As previously mentioned, MLM techniques provide

substantial incremental utility over rmANOVA models by allowing the testing of naturally nested data (for our purposes, within-subject independent variables) in conjunction with continuous outcome variables, and do not have the drawbacks inherent to rmANOVA. We will first describe overall model characteristics and how they were estimated, and then will provide a specific outline of our model fitting, permutations on the model, and hypothesis testing strategy.

Model characteristics and parameters.

All fitted models were two level MLMs, with level 1 referred to as the “Task” level and level 2 referred to as the “Person” level, as each Task variable is nested within each Person (i.e., we measured each participant on each trial type on the PIG task). Accordingly, all intra-individual difference (i.e., within-subject) variables pertaining to the PIG (stimulus class, experimental outcome measure when used as a predictor) were entered at level 1, and all inter-individual difference (i.e., between-subjects) variables (personality variables, demographic variables) were entered at level 2. The dependent variable is by definition a level 1 variable in MLM. Unless otherwise noted, stimulus dimension (“Stimulus”) was entered as a continuous variable (consistent with studies identifying generalization as an inherently continuous process; e.g., Tenenbaum & Griffiths, 2001) in which Δ CS- was considered the control condition (i.e., no Stimulus effect) and therefore was coded as 0, the oCS- was coded as 1, GS1 as 2, GS2 as 3, GS3 as 4, and CS+ as 5, resulting in a continuous 0-5 scale.

Pairwise deletion was used to exclude missing outcome data for each participant as needed; MLM assumptions are not violated if cases with partially missing data are present (Gelman & Hill, 2006). All dependent variables were unstandardized per

regression best-practices (Fischer & Milfont, 2010), as well as due to particular complications that arise with this technique if done in MLMs (Moeller, 2015). Centering/standardization is typically accomplished through two methods in MLM: centering at the grand mean (CGM) and centering within cluster¹⁶ (CWC), which can also be thought of as centering using the grand mean of all participants compared with centering repeated measures within each participant using their individual mean, respectively (Enders & Tofighi, 2007). Each method of centering/standardization is used for different reasons related to *a priori* interest in level 1 effects, level 2 effects, or cross-level interactions (i.e., interactions between one or more level 1 and level 2 variables). Unless otherwise noted, we used CGM for level 2 variables. This decision was made based on recommendations (Enders & Tofighi, 2007) that 1) CGM should be used when testing level 2 variables while controlling for level 1 variables (e.g., effect of dispositional trait while controlling for stimulus type to quantify main effect); and 2) CGM is recommended for interactions between level 2 variables (i.e., using interactions between two dispositional traits as predictors). Our models did not contain level 1 variables that are appropriately transformed with a centering technique (i.e., they are variables representing the stimulus dimension as defined by the PIG task and are a constant across participants) with the notable exception of APIC analyses; this exception is discussed in the following section. Further, scaling using z-scores was the specific centering technique used for CGM; we made this decision in accordance with

¹⁶ The “clustering” terminology is reflective of a strong MLM tradition within the educational sciences and thus many MLM examples and terms refer to clusters of people at level 1 (i.e., children clustered within a class) – for our purposes, clustering refers to clusters of task data, not clusters of people. In other words, each person is a cluster.

recommendations for improving interpretability of interaction terms for two or more continuous variables (Aiken, West, & Reno, 1991; Friedrich, 1982). Related to this, standardized interaction terms were defined using the product of standardized lower-order terms, again in accordance with standard recommendations (Aiken et al., 1991; Friedrich, 1982). For predicting an assumed normally-distributed outcome (i.e., startle and risk ratings), standard MLM (i.e., linear mixed modeling [LMM]) was used to fit the data. For predicting a binomially-distributed outcome (i.e., avoidance), the generalized extension of a standard MLM was used (i.e., generalized linear mixed model; GLMM) to fit the data, with predictors linked to the binomial outcome via a logit-link function (as in logistic regression). All model parameters were estimated using maximum likelihood estimation (MLE), which is standard for MLMs containing a sufficient number of level 2 observations/sample size (Gelman & Hill, 2006; Maas & Hox, 2005), and the “optimx” family of optimization algorithms were used as the computational method for all models to maximize the likelihood function central to MLE (Nash, 2014). Two forms of R^2 were obtained for each MLM using the sjstats package (Lüdtke, 2018): marginal R^2 , which is an estimate of variance accounted for by fixed-effects only, and conditional R^2 , which is an estimate of variance accounted for by fixed-effects plus random-effects (i.e., the entire model) (Nakagawa, Johnson, & Schielzeth, 2017). For all models, we visually inspected residual plots and qq-plots to detect obvious deviations from homoscedasticity or normality (Gelman & Hill, 2006). Finally, it should be noted that due to the complexity of the analyses and the relatively small number of observations for each participant (compared with the bulk of studies using MLM; Kondo et al., 2009; Sellström & Bremberg, 2006), it is possible that a small number of the more complex models will not

be successfully fit (i.e., model convergence will not be possible and no parameters can be estimated); this will be addressed on a case-by-case basis.

Initial model fitting plan.

Standard MLM best practices in psychological research involves the fitting of successive models in a hierarchical fashion (i.e., each successive model contains all terms from the previous model and adds additional terms) with the goal of finding the optimal balance between model parsimony and explanatory power while also building a model that is useful for addressing the scientific questions under investigation (Aguinis, Gottfredson, & Culpepper, 2013; Barr, Levy, Scheepers, & Tily, 2013; Bolker et al., 2009; Gelman & Hill, 2006). As discussed previously, this approach has been strongly recommended for primary analyses in experimental fear generalization studies (Vanbrabant et al., 2015) and has been adopted by multiple generalization research groups for their studies (e.g., FeldmanHall et al., 2018; Ginat-Frolich, Klein, Katz, & Shechner, 2017; Lenaert, van de Ven, Kaas, & Vlaeyen, 2016; Lommen et al., 2017).

Given we have three dependent variables, we conducted three parallel “lines” of analyses. For each line, we initially started with a random-intercept only model that did not contain any fixed effects. These models, sometimes referred to as “null models”, provide the rationale for conducting MLM in the first place, as they allow us to answer if there is sufficient within-subjects variance (i.e., dependencies) that needs to be accounted for through a MLM technique (Gelman & Hill, 2006). An intraclass coefficient (ICC), which typically ranges from 0 to 1, is used to quantify the proportion of within-subject variance (how similar observations are within a person compared to similarity between people), with higher ICCs indicative of a higher proportion of within-subjects variance

and need for MLM (Koo & Li, 2016; Musca et al., 2011). We calculated adjusted ICC for models that include a random-slope term due to concerns about an unadjusted ICC not reflecting the correct variance proportions in these models (Goldstein, Browne, & Rasbash, 2002; P. C. D. Johnson, 2014).

If the ICC confirmed the need for MLM, we then followed the general approach outlined by Vanbrabant et al. (2015). “Base models” (models without dispositional traits of interest) were constructed, starting with a simple random-intercept only model (Model 0 in our framework) that allows each person to have their own intercept and also contain the continuous “Stimulus” variable at level 1 as a fixed effect and Gender (coded dichotomously) entered as a level 2 fixed effect. We included Gender as a covariate based on 1) preliminary analyses from this and a related sample using the same task that significant gender effects exist for multiple dependent variables, most notably avoidance, and 2) ongoing research documenting gender effects in experimental studies of fear and anxiety (McLean & Anderson, 2009), including fear conditioning studies (e.g., Rosenbaum et al., 2015).

Next, we added a Stimulus random-effect parameter, this parameter estimates a random slope for each participant (i.e., the line of best fit given the values observed for each level of Stimulus for that particular person), for our first random-intercept/random-slope model (Model 1). The fixed-effect term for Stimulus is maintained from the previous model, as is always the case when adding a random-effect. In the next model (Model 2), we add a quadratic effect of Stimulus to our model (Stimulus²) by

exponentiating the continuous variable and including it as a fixed-effect in our model¹⁷; this is both in line with recommendations from Vanbrabant et al. (2015) and prior generalization findings of stronger quadratic components indexing less generalization (for review, see Dymond, Dunsmoor, et al., 2014). In the final model (Model 3), we also add a Stimulus² random-effect, which allows us to estimate both linear and quadratic parameters for each participant's individual slope. To ascertain which base model was optimal for further testing, we used the “keep it maximal” data-driven approach recommended by Barr et al. (2013), in which the model with the most complex random-effects structure that is allowed by the data and study design considerations. In our study, this was operationalized as 1) the most complex model in terms of random-effects structure; 2) the model contained at least one random-slope term (as the goal of the study is to identify individual differences in generalization, which suggests individual slope coefficients are needed); 3) model fitting did not result in a singular fit (i.e., one of the random-effects accounts close to or exactly 0/the model is overfitted); and 4) the model was a significantly better fit for the data than the previous model in the series (i.e., a model with exactly one less term), indicating that the additional term was significantly improving the predictive power of the model while the model also remained relatively parsimonious. To determine this last criterion, we used log-likelihood tests (Bolker et al., 2009) to assess differences in model fit (see the later section on hypothesis testing for a more thorough description of this test).

Model fitting plan for modeling effects of interest.

¹⁷ Note that although higher-order polynomials (e.g., cubic) are potentially informative in this context, there is a high risk of overfitting the data given the task design (Vanbrabant et al., 2015) and thus we do not go beyond the quadratic polynomial

Once a base model was identified as meeting our established criteria, we built a series of models predicated on published recommendations (Gazendam et al., 2014; Harrison et al., 2018; Vanbrabant et al., 2015), the current investigation's scientific aims, and our hypothesis testing plan. These series of models were defined in a hierarchical fashion and, an example of which can be seen in Appendix A. We will use two an example involving BFAS traits to provide a detailed illustration of our procedure, one that illustrates our procedure for testing a single individual difference measure, multiple individual difference measures, and interactions among those measures.

For the first example: For models testing Neuroticism, we first entered Neuroticism in the base model as a level 2 fixed-effect. This model now outputs a Neuroticism coefficient and allows for testing of a “main effect” of Neuroticism. Next, we interacted Neuroticism with the Stimulus (linear only) variable, and in a separate model interacted Neuroticism with both the Stimulus and Stimulus² (quadratic) variable, which results in two separate two-way interactions in the same model that together describe a curvilinear relationship. The former model provides an estimate of the effect of Neuroticism on the dependent variable that is conditional on the level of Stimulus when modeled linearly with no curve, the latter model provides an estimate of the effect of Neuroticism on the dependent variable that is conditional on the level of Stimulus when also modeled linearly and but now is allowed to curve (i.e., has a quadratic component). Alternatively, this can be viewed as testing if Neuroticism helps explain variability in the estimated generalization gradient when modeled only using linear components or with a polynomial (quadratic) component.

After this stage, further models would be constructed if the candidate personality dimension was a higher-order trait that was comprised of two or more lower-order dimensions (e.g., aspect or facet) and was specified as a candidate lower-order dimension in our *a priori* hypotheses. For this example, two aspects underlie Neuroticism (Withdrawal and Volatility) and we have predictions related to both, thus additional models are built to test these aspects. A sub-series of models are built using the exact technique as described in the previous paragraph, one for Withdrawal and, separately, one for Volatility. Additionally, a third series of models were built that is similar but contains both aspects in the same models. In the initial model, Withdrawal and Volatility are modeled as separate level 2 fixed-effects, providing “main effects” for one aspect while controlling for the other. The next pair of models then contains cross-level interactions, with Volatility and Withdrawal separately interacted with 1) the Stimulus variable or 2) both the Stimulus and Stimulus² (quadratic) variable; this latter model results in two pairs of separate two-way interactions in the same model. Next, depending on specific hypotheses, an additional model contains a within-level 2 interaction between the two aspects, which allows for testing of conditional effects of one aspect on the other that might be obscured when only modeling the higher-order trait (which is mathematically the average of both aspects). Relatedly and also depending on hypotheses, two subsequent models would then interact this within-level 2 two-way interaction term with one of the two Stimulus terms to create two separate cross-level three-way interaction models. Additionally, due to potential concerns about violations of model assumptions driven by multicollinearity in the more complex models (Echambadi & Hess, 2007; Gelman & Hill, 2006; Snijders & Berkhof, 2008), descriptive diagnostics,

most pertinently variable inflation factor (VIF) were produced and examined for all models containing more than one level-2 trait variable.

Modifications in model fitting plan for APIC-G/APIC-CS+ analyses.

An exception to the previously outlined procedures is our plan for APIC-G/APIC-CS+ analyses. To facilitate testing of these relationships and to approximate prior research using models without random-effects (van Meurs et al., 2014), we made a number of important modifications. First, there were six “lines” of models. There were two for APIC overall (not distinguishing between APIC-G and CS+); two for APIC-G and two for APIC-CS+: within each type of APIC analysis, there was one model with startle values as a level-1 predictor, and one with risk rating values as a level-1 predictor, with each Pavlovian value yoked to one avoidance value based on how our data was structured. Second, to directly model APIC, initial model fitting included the Pavlovian fear index at level-1, and all models interacted the Pavlovian level-1 value with one or both Stimulus dimensional variables. Third, due to the nature of the PIG design used in this study, we were not able to model each avoidance observation separately, as risk ratings are only assessed three times per trial type during generalization compared with six times per trial type for avoidance. Due to this mismatch, we used the average response on each trial type for all Pavlovian predictors and for avoidance in these models (e.g., average startle, risk rating, and avoidance to CS+, GS3, GS2, etc.). Fourth, given that APIC modeled using MLM involves a level-1 predictor (Pavlovian x Instrumental, both are level-1 variables), Pavlovian variables were centered using CWC (in our study, referred to as within-subject centering) per standard guidelines (Enders & Tofighi, 2007). This removes between-subjects variation for level-1 predictors (Raudenbush & Bryk,

2002) and therefore accounts for the dependency between the two Pavlovian variables and avoidance decisions, which is by definition a non-independent/within-subjects relationship because all three variables are collected during the PIG task for all participants. Fifth, the outcome variable, avoidance, is typically modeled as binary outcome in “success/failure” form (i.e., 0 or 1). However, for these analyses, avoidance was modeled as the proportion of avoidance decisions (# of avoidance decisions/total trials). As the proportion is still bound by 0 and 1, it is permissible, if not ideal, to use this form of our avoidance data in the logistic regression framework. At this point, the overall APIC model did receive any additional modifications. However, one additional change was made for APIC-G and APIC-CS+ models: the coding for Stimulus dimension was modified in accordance with the effect of interest. For both APIC-G and APIC-CS+ the Δ CS- was kept as the referent and coded as zero; for APIC-G the GS1 was coded as 1, the GS2 as 2, and the GS3 as 3 (oCS- and CS+ excluded; and for APIC-CS+ only the CS+ was added in addition to the Δ CS- and coded with a 1 (i.e., CS+ and Δ CS- were dummy coded).

Hypothesis testing plan.

We used two different techniques for hypothesis testing. The type of test depended on the specific hypothesis that we were addressing. For predictions related to specific main effects or interactions (and their corresponding coefficients), the “lmerTest” package was used (Kuznetsova, Brockhoff, & Christensen, 2017) to provide appropriate degrees of freedom for hypothesis testing of fixed-effect predictors. In the case of standard linear MLMs (i.e., LMMs), which we used for models predicting startle and risk rating outcomes, we used Wald F-tests in conjunction with Satterthwaite's

method (Satterthwaite, 1946), which provides an effective degrees of freedom for data that has an overall distribution comprised of multiple independent sources of variation (in our MLMs, this is due to each participant having their own within-task variance) and has been proposed as an acceptable technique for testing MLM components (e.g., Berkhof & Snijders, 2001). The Satterthwaite estimated degrees of freedom are used in standard F-tests for each coefficient of interest, and can be interpreted in a similar fashion to how F- or t-tests for standard multiple regression models/individual coefficients are interpreted. For generalized MLMs (i.e., GLMMs), which we used for models predicting avoidance, we again used a Wald's test; in this case the Wald chi-square goodness-of-fit test that yields a χ^2 statistic. This test indicates if a model including the tested predictor outperforms a model that does not include the tested predictor (Hosmer Jr, Lemeshow, & Sturdivant, 2013), and in terms of precision and bias performs similarly to a t-test in larger samples (Agresti & Kateri, 2011; Hauck Jr & Donner, 1977). Given that it is a goodness-of-fit test, there is one degree of freedom, and the reported N for each test refers to number of analyzed Level 1 observations (i.e., number of individual avoidance decisions available for analysis), not Level 2 observations (i.e., number of participants). Both forms of Wald tests previously described are commonly used for testing fixed-effects in MLM (Berkhof & Snijders, 2001; Maas & Hox, 2005). These tests were used for Aims 1, 3, and 4, which all centered on testing specific MLM variance components.

The second type of test we used was the log-likelihood ratio test (LRT), another goodness-of-fit test. This test was used at the level of the whole model, and specifically to determine if one model is a significantly better fit for the data compared with a different model (i.e., better predictive power without adding excessive error to the model)

and is a standard technique for comparing models in an MLM framework (Gelman & Hill, 2006). LRTs were performed to either compare initial models that did not include a predictor of interest (e.g., personality variable) to a model that contains the predictor of interest, or to hierarchically compare goodness-of-fit across a series of models that include predictor(s) of interest (e.g., comparing the model that only contains Neuroticism to the model that contains the Neuroticism aspects, or to the model that contains Neuroticism, Extraversion, and Conscientiousness). The LRTs reported in this dissertation were performed on the deviance statistic that is calculated for all MLMs in lme4; this statistic is a numerical index of how well a nested model fits compared with a hypothetically perfectly fitted model, with a smaller deviance statistic corresponding with a better fitting model. The LRT compares the likelihood of obtaining a “true” deviance statistic using the chi-square distribution, and therefore we report χ^2 statistics for these tests. We also provide the marginal R^2 (denoted herein as R_M^2) when reporting LRT results to provide an effect size estimate (i.e., magnitude of prediction improvement). Deviation scores, as well as another index of model fit (Akaike Index Criteria [AIC]) are provided for models in their corresponding tables. These tests were used for Aim 2 only.

It should also be noted that although many more models are created in the current study’s approach than with rmANOVA, the number of hypothesis tests are comparable, as rmANOVA main effects and lower-order interactions can largely be interpreted without regard to how one affects the other (e.g., a main effect test in rmANOVA is still interpreted as a main effect even if there is an interaction including the main effect factor in the same model) and therefore multiple hypothesis tests can be done using a single model. In contrast, there are many more models constructed in this study than in

equivalent studies using rmANOVA techniques, but fewer hypothesis tests per model (e.g., we do not interpret lower-order terms in models that contain a two-way interaction).

Interaction follow-up testing plan.

Our plan for follow-up analyses on significant interactions (as indicted by the χ^2 and F-tests outlined in the previous section) involved both quantitative and graphical techniques. First, it should be noted that only the highest-order interaction in a model was interpreted and subject to follow-up analyses unless otherwise noted. This is because interpretation of lower-order interactions in models with multiple levels of interactive effects is not recommended, especially if intending to interpret as an unconditional interaction (i.e., as if the higher-order interaction wasn't in the model) (A. F. Hayes, 2018; A. F. Hayes, Glynn, & Huge, 2012). This means that for models containing three-way interactions, we did not interpret coefficients for two-way interactions in the same model.

For two-way interactions that are equivalent to generalization gradients from prior work (i.e., a Stimulus dimension interacted with an individual difference), we generated two types of plots. The first plot depicts raw dependent variable values (i.e., actual values) on the y-axis graphed across the categorical Stimulus continuum on the x-axis, and the individual difference variable dichotomized (for graphing purposes only) into averaged high (above the median) and low (below the median) groups and represented with separate lines depicting point estimates at each Stimulus level. These plots resemble traditional generalization gradients and therefore both the overall shape (e.g., degree of departure of the slope from linearity) and specific slope segments (e.g., relative steepness or shallowness of the segment) are interpreted. The second plot involves fitted values

from the model containing the significant interaction, again plotted across the Stimulus dimension and for high/low averaged levels of the individual difference, but also with 95% confidence intervals visualized for each plotted line. These plots accompany the generalization gradients for more precise interpretation of results in the context of an interaction with two continuous variables, which are what we test in the current dissertation's multilevel models, as opposed to the two categorical variable interactions typically seen in the generalization literature using rmANOVA (e.g., anxiety disorder group x categorical stimulus dimension). More specifically, these fitted line plots depict the strength of the linear relationship between the continuous individual difference variable and the dependent variable as a function of change in stimulus value, with larger gaps between the plotted lines and their confidence intervals indicating stronger relations between the individual difference and the dependent variable. The traditional generalization gradients do not provide this level of nuance and ability to visualize linear change across the stimulus continuum due to the use of actual values (as opposed to fitted values) and relying on point estimates at each stimulus level. Of note, these fitted values plots are only useful to interpret generalization effects when used in conjunction with the generalization gradients, as the shape of the gradient is largely lost when a fitted line is visualized instead of actual values.

For three-way interactions (i.e., those resulting from APIC analyses, which contain Trait x Stimulus x Pavlovian variable interactions), we used a simple-slopes approach, modified to account for multilevel modeling techniques (Aiken et al., 1991; Bauer & Curran, 2005; Hox, Moerbeek, & Van de Schoot, 2017), to decompose the interaction. Specifically, after identifying the significant interaction, the moderator of

interest was re-centered so that its new center was 1 standard deviation above the actual mean (i.e., 1 SD is added as a constant to all z-scored values) and the model was re-run. This was repeated for 1 standard deviation below the mean. These two additional models now contained a lower-order term with the non-moderating interaction variable that yields a simple slope coefficient which could be tested to determine if the high (+1SD) and/or low (-1SD) moderator slope is significant and contributing to the original significant interaction (as this lower-order term now represents the effect of the non-moderating interactive variable at the zero-point of the re-centered moderator, which means the -1SD model provides a high zero-point/high simple slope and the +1SD model provides the low zero-point/low simple slope). The next step was visualization of the interaction, again using multiple types of plots to visualize the interaction at focal values for multiple predictors. The first type of plot was similar to the first type of plot used for two-way interactions, except that fitted values were used, the y-axis corresponded to avoidance (represented as a percentage), the x-axis corresponded to the Pavlovian response variable (startle or risk ratings), and separate plots for each stimulus level were generated. The individual difference variable (the moderator in APIC models) was again dichotomized into high/low groups based on the median. These plots included both a scatterplot component, with each point colored according to group designation, and two summary fit lines that were again coded by group. This combination provides both an estimation of the potential moderation effect at each stimulus level, as well as information about the overall Pavlovian x avoidance relationship (which can be seen in the overall pattern of the plotted points. Additionally, we generated separate plots visualizing the simple regression slopes described previously (each plot contained the +1

and -1 SD slopes) at the level of CS+ and GSs (i.e., the stimuli levels of interest for APIC). These simple slope plots allowed for interpretation of relative effect of the moderator on avoidance, as opposed to the absolute effect seen in the fitted value scatter plots (e.g., a potentially large moderating effect for the GS1 might be obscured in the scatterplot due to overall smaller range of avoidance values associated with the GS1). It should also be noted that we refrained from plotting the simple slope at the level of the mean, which is common for simple slopes analyses, because 1) it simplified visualization; 2) only two values (e.g., high/low values) are needed to interpret a linear continuous moderator; and 3) there is not a unique meaning to an average level of most of the personality variables tested in this study. Both types of plots were interpreted to decompose three-way interactions, with the simple slopes plots prioritized if the summary fit lines or individual points in the scatterplots were ambiguous (e.g., difficult to distinguish between the lines or clusters of points) or if a visible outlier was present, but did not merit exclusion due to a lack of statistical evidence that the particular point was an outlier and unduly influenced results. It should also be noted that the scatterplots represent a more modest split of the continuous individual difference variable, whereas the simple slopes represent relatively extreme high/low values, which can lead to slight differences between the plots. Despite this, we contend that both types of plots were necessary to obtain an adequate visual summary of our results.

Results

Preliminary Analyses

Manipulation checks.

Pre-acquisition.

In line with previous generalization work and our group's previous study using the PIG, we did not find significant within-subject differences in startle ($p = .41$, *ns*) or risk ratings ($p = .23$, *ns*) at pre-acquisition (i.e., prior to conditioning).

Acquisition.

See Figure 4 for graphed results. There was a significant main effect of Trial-Type, $F(2, 279) = 97.62$, $p < .001$, $\omega^2 = .158$, for startle data, reflecting potentiated startle to the CS+ compared with the oCS-, $t(279) = 10.85$, $p < .001$, and with the Δ CS-, $t(279) = 11.71$, $p < .001$; the difference between the oCS- and Δ CS- was not significant, $t(279) = 1.01$, $p = .31$.

There was also a significant main effect of Trial-Type, $F(2, 283) = 1326.96$, $p < .001$, $\omega^2 = .728$ for risk rating data, reflecting higher ratings for the CS+ compared with the oCS-, $t(283) = 38.37$, $p < .001$, and with the Δ CS-, $t(283) = 43.36$, $p < .001$. The oCS- was also rated significantly higher than the Δ CS-, $t(283) = 2.85$, $p = .004$.

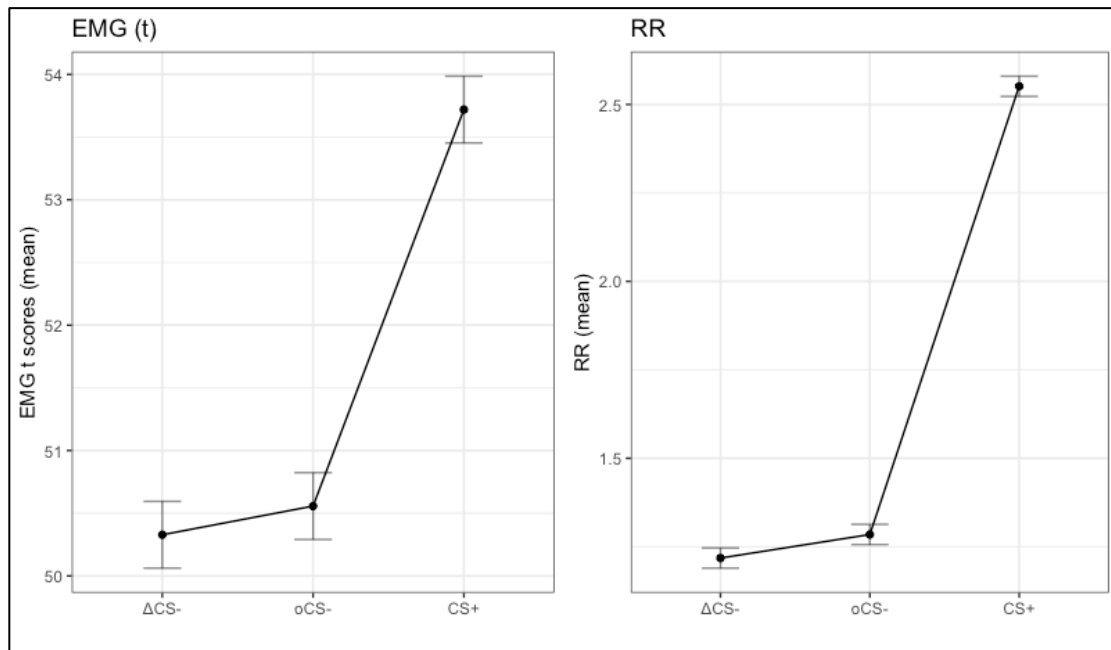


Figure 4. Average startle (left) and risk rating (right) responses to the triangular CS-, circular CS-, and CS+ during the acquisition phase. Startle responses are represented in t-

score form (within-subject standardization). Error bars represent standard error of the mean. CS- = conditioned safety cue; CS+ = conditioned danger cue; RR = risk rating.

Generalization.

Startle.

Potentiated startle to the CS+ relative to the safety stimuli persisted into the generalization phase (oCS-: $t[287] = 16.88, p < .001$; Δ CS-: $t[287] = 16.94, p < .001$). There was a main effect of Trial-Type, $F(5, 287) = 315.73, p < .001, \omega^2 = .515$, which was consistent with generalization across the stimulus continuum. A significant quadratic trend, $F(1, 287) = 280.61, p < .001, \omega^2 = .475$, confirmed a generalization effect in the startle data. There continued to be no significant difference in startle potentiation between the oCS- and Δ CS-, $t(287) = -0.6394, p = .52$). See Figure 5 for graphed generalization gradient.

Risk ratings.

Elevated risk ratings to the CS+ relative to the safety stimuli persisted into the generalization phase (oCS-: $t[287] = 55.51, p < .001$; Δ CS-: $t[287] = 54.59, p < .001$). There was a main effect of Trial-Type, $F(5, 287) = 1141.92, p < .001, \omega^2 = .706$, which was consistent with generalization across the stimulus continuum. A significant quadratic trend, $F(1, 287) = 732.64, p < .001, \omega^2 = .570$, confirmed a generalization effect in the risk rating data. The significantly higher risk ratings to the oCS- compared with the Δ CS- during acquisition did not persist into generalization, $t(287) = -1.43, p = .15$). See Figure 5 for graphed generalization gradient.

Avoidance.

Rate of avoidance for the CS+ was significantly higher compared with the safety stimuli (oCS-: $t[287] = 26.34, p < .001$; Δ CS-: $t[287] = 26.35, p < .001$), indicating

successful instrumental conditioning. There was a main effect of Trial-Type, $F(5, 287) = 504.48$, $p < .001$, $\omega^2 = .497$, which was consistent with generalization across the stimulus continuum. A significant quadratic trend, $F(1, 287) = 441.16$, $p < .001$, $\omega^2 = .407$, confirmed a generalization effect in the avoidance data. We did not find a significant difference in avoidance rate between the oCS- and Δ CS-, $t(287) = -0.188$, $p = .85$). See Figure 5 for graphed generalization gradient.

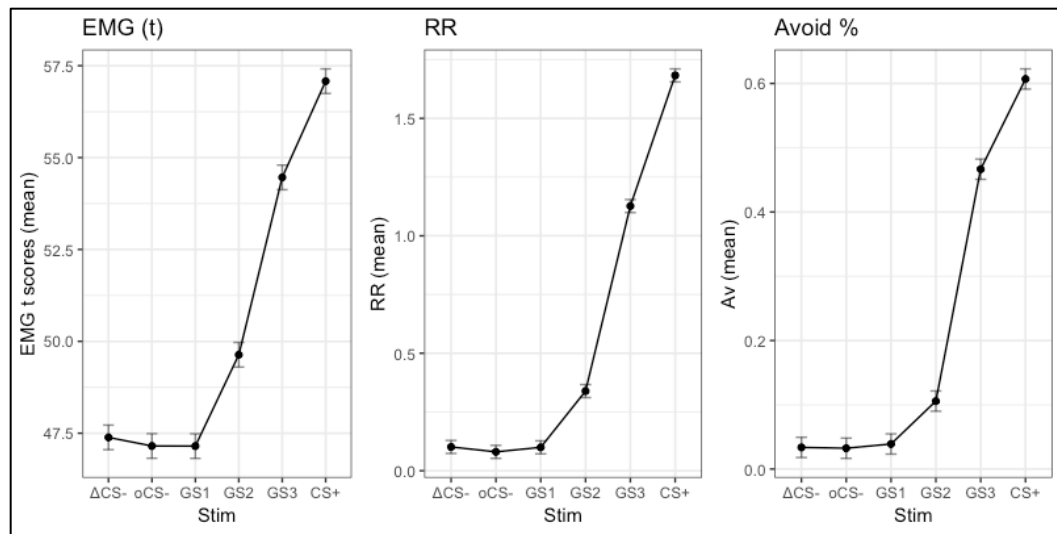


Figure 5. Average startle (left), risk rating (center), and avoidance (right) responses to the triangular CS-, circular CS-, GSs, and CS+ during the generalization phase. Startle responses are represented in psychometric t-score form (within-subject standardization). Avoidance responses are represented in percentage, with 100% equal to always avoiding. Error bars represent standard error of the mean. Av = avoidance; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimuli; RR = risk rating.

Interim summary of manipulation checks.

All manipulation checks were successful, and indicated that across all indices 1) participants did not have a differential response to the PIG stimuli prior to conditioning; 2) learned the CS+/US contingency; 3) demonstrated discrimination between the CS+ and the safety stimuli and that this persisted from the acquisition phase to the generalization phase; 4) that the task successfully elicited generalization from the CS+ to

the GSs; and 5) that participants learned the instrumental association. Additionally, no differences were found between the two conditioned safety cues (oCS- and Δ CS-), with the exception of participants rating the oCS- as significantly more dangerous than the Δ CS- within the acquisition phase only; this effect did not persist into the generalization phase. These results are close to identical to those found in the initial study using the PIG paradigm (van Meurs et al., 2014), with the exception that the prior study found a trend towards greater startle to the oCS- compared with the Δ CS- during acquisition, and did not find a risk rating difference during this phase.

Participant characteristics.

For descriptive statistics for scales analyzed in the current study, see Table 3. For an expanded set of tables that includes descriptive statistics for scales from the measures administered in this study, including those not used for analyses (e.g., Agreeableness and Openness/Intellect from BFAS, individual facets from PID-5) see Appendix B.

Table 3. Descriptive statistics for broadband (BFAS, PID, PB) and narrowband measures (TF-44, IUSF, ASI, STAI-T).

Variable	Mean	SD	Min	Median	Max
<u>BFAS (trait)</u>					
Neuroticism	2.7	0.72	1.15	2.67	4.6
Conscientiousness	3.49	0.57	1.35	3.5	4.75
Extraversion	3.51	0.64	1.7	3.52	4.8
<u>BFAS (aspect)</u>					
(N) Withdrawal	2.82	0.75	1.1	2.8	4.6
(N) Volatility	2.57	0.82	1	2.6	4.9
(C) Industriousness	3.46	0.67	1.2	3.5	5
(C) Orderliness	3.53	0.64	1.5	3.6	4.8
(E) Assertiveness	3.38	0.71	1.6	3.4	5
(E) Enthusiasm	3.65	0.77	1.1	3.7	5

PID-5 (trait)

Negative Affectivity	0.9	0.42	0.12	0.86	2.11
Disinhibition	1.01	0.37	0.28	0.96	2.33
Detachment	0.74	0.44	0.04	0.67	2.17

PID-5/BFAS (PB) composite

Distress-PB	0	0.81	-1.34	-0.08	2.43
Industriousness-PB	0	0.73	-2.88	0.13	1.35
Orderliness-PB	0	0.87	-2.17	-0.01	2.3
Assertiveness-PB	0	0.8	-1.93	0.05	2.04
Enthusiasm-PB	0	0.75	-2.39	0.08	1.32

Narrowband traits

TF-44	62.08	21.78	0	62	122
IUSF	31.07	8.99	12	31	59
ASI	22.4	12.15	0	20	59
STAI-T	41.95	9.94	24	41	70

Notes: BFAS aspect labels are preceded by the corresponding higher-order trait.

Component scales of PB variables are standardized prior to calculation, therefore means for all PB scales are zero. ASI = Anxiety Sensitivity Index; BFAS = Big Five Aspect Scale; C = Conscientiousness; E = Extraversion; IUSF = Intolerance of Uncertainty – Short Form; PB = PID-5/BFAS composite; PID-5 = Personality Inventory for DSM-5; N = Neuroticism; STAI-T = State-Trait Anxiety Inventory – Trait version; TF-44 = Trait Fear – 44 item.

Trait individual difference variables.

Table 4 contains correlations between all trait variables that were included in main analysis models. As can be seen, the majority of the trait variables significant correlated with the other trait variables measure, and therefore the magnitude of correlation is the more important information presented in Table 4.

Task variables.

Table 5 contains correlations between all variables (startle, risk ratings, and avoidance) at each stimulus level for the generalization phase. As can be seen, most avoidance and risk rating indices significantly positively correlated with each other,

whereas raw startle was not related to the other two task variables when the association was measured with simple Pearson correlations.

Table 4. Trait variable correlations.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. DE	-																			
2. DI	.39*	-																		
3. NA	.76*	.59*	-																	
4. W	.51*	.31*	.69*	-																
5. V	.31*	.39*	.63*	.66*	-															
6. In	-.39*	-.47*	-.46*	-.52*	-.43*	-														
7. Or	-.15*	-.11	-.06	-.05	0	.50*	-													
8. En	-.72*	-.12*	-.39*	-.37*	-.18*	.34*	.23*	-												
9. As	-.33*	.05	-.19*	-.36*	-.07	.36*	.23*	.49*	-											
10. N	.44*	.38*	.72*	.90*	.92*	-.51*	-.03	-.30*	-.23*	-										
11. C	-.32*	-.34*	-.31*	-.34*	-.25*	.87*	.86*	.33*	.34*	-.32*	-									
12. E	-.61*	-.05	-.34*	-.42*	-.15*	.41*	.26*	.87*	.85*	-.30*	.39*	-								
13. DPB	.64*	.42*	.90*	.87*	.65*	-.49*	-.05	-.33*	-.27*	.82*	-.32*	-.35*	-							
14. IPB	-.41*	-.86*	-.55*	-.37*	-.39*	.76*	.43*	.19*	.13*	-.42*	.69*	.19*	-.45*	-						
15. OPB	.09	.23*	.22*	.11	.19*	.31*	.87*	.05	.19*	.17*	.67*	.14*	.16*	.19*	-					
16. APB	-.35*	-.12	-.38*	-.39*	-.13*	.34*	.14*	.34*	.80*	-.28*	.28*	.65*	-.39*	.21*	.05	-				
17. EPB	-.95*	-.32*	-.64*	-.41*	-.23*	.33*	.17*	.83*	.34*	-.35*	.29*	.68*	-.49*	.33*	-.06	.30*	-			
18. STAIT	.60*	.33*	.73*	.80*	.58*	-.46*	-.06	-.36*	-.28*	.75*	-.31*	-.37*	.84*	-.39*	.1	-.33*	-.47*	-		
19. IUSF	.42*	.19*	.52*	.48*	.41*	-.17*	.22*	-.26*	-.14*	.48*	.02	-.23*	.51*	-.11	.36*	-.22*	-.38*	.48*	-	
20. ASI	.34*	.29*	.53*	.51*	.43*	-.27*	0	-.16*	-.03	.51*	-.16*	-.11	.55*	-.24*	.21*	-.11	-.29*	.50*	.49*	-
21. TF44	.34*	-.17*	.38*	.61*	.35*	-.30*	.04	-.36*	-.55*	.52*	-.16*	-.52*	.53*	.04	.07	-.49*	-.32*	.54*	.51*	.38*

Note: As = Assertiveness; APB = Assertiveness PID-BFAS; ASI = Anxiety Sensitivity Index; C = Conscientiousness; DE = Detachment; DI = Disinhibition; DPB = Distress PID-BFAS; En = Enthusiasm; EPB = Enthusiasm PID-BFAS; E = Extraversion; In = Industriousness; IPB = Industriousness PID-BFAS; IUSF = Intolerance of Uncertainty Short Form; OPB = Orderliness PID-BFAS; Or = Orderliness; N = Neuroticism; NA= Negative Affectivity; STAIT = State Trait Anxiety Inventory – Trait version; TF-44 = Trait Fear – 44 item; W = Withdrawal; V = Volatility.

* $p < .05$

Table 5. Task variable correlations.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 Startle: CS-	-																	
2 Startle: oCS-	.95*	-																
3 Startle: GS1	.94*	.94*	-															
4 Startle: GS2	.91*	.95*	.94*	-														
5 Startle: GS3	.83*	.87*	.87*	.91*	-													
6 Startle: CS+	.81*	.84*	.84*	.87*	.95*	-												
7 RR: CS-	-.06	-.06	-.05	-.06	-.08	-.09	-											
8 RR: oCS-	-.08	-.07	-.04	-.05	-.08	-.11	.39*	-										
9 RR: GS1	-.07	-.06	-.01	-.04	-.08	-.08	.37*	.68*	-									
10 RR: GS2	-.04	-.02	.02	.02	0	-.03	.31*	.47*	.53*	-								
11 RR: GS3	0	.03	.07	.08	.09	.04	.12*	.19*	.17*	.48*	-							
12 RR: CS+	.03	.02	.08	.07	.08	.06	.14*	.14*	.14*	.26*	.41*	-						
13 Av: CS-	.01	-.01	.01	-.01	-.03	-.02	.45*	.21*	.18*	.22*	.13*	.1	-					
14 Av: oCS-	-.01	0	.02	-.01	-.03	-.04	.26*	.34*	.41*	.30*	.08	.09	.46*	-				
15 Av: GS1	-.02	0	.03	0	.02	-.01	.35*	.37*	.49*	.39*	.14*	.09	.46*	.73*	-			
16 Av: GS2	.01	0	.06	.05	.05	.02	.24*	.32*	.36*	.48*	.34*	.21*	.47*	.55*	.69*	-		
17 Av: GS3	0	0	.05	.04	.07	.06	.08	.12*	.15*	.23*	.40*	.28*	.22*	.20*	.30*	.48*	-	
18 Av: CS+	.05	0	.07	.05	.09	.07	.05	.08	.1	.12*	.25*	.32*	.15*	.15*	.21*	.34*	.77*	-

Note: Av = Avoidance; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; RR = risk ratings.

* $p < .05$

Main analyses

Results are organized by the stated aims of this dissertation, and then further organized by dependent variable and type of personality variable. For aims involving tests of specific effects (i.e., individual model coefficients), results are further organized by the specific personality trait. To maintain a semblance of brevity, we only report full statistics for significant hypothesis tests. Full model and hypothesis statistics for all models are available on request from the author. We report standardized coefficient values (β) and 95% confidence intervals (CI) for each test when reporting results for individual effects (i.e., variance components). For generalized MLMs, we report β s in odds ratio (OR) form (for interpreting lower-order coefficients, OR values > 1 indicating increasingly greater chance of avoidance for each unit increase in the predictor, values < 1 indicating an increasingly lower chance of avoidance for each unit increase in the predictor, and with OR = 1 equal chance; this interpretation changes for interaction coefficients).

Aim 1 (A1): Identifying personality predictors of generalization.

A1: Startle.

Testing of the null-model indicated a multilevel model is appropriate for these data (ICC = .67). Initial model fitting resulted in Model 2 (random-intercept, random-slope for Stimulus dimension, fixed-effect for Stimulus² dimension) as the most acceptable model for the following startle analyses based on our established criteria, with both the random-intercept and all Stimulus effects contributing significantly to the model and adjusted ICC = 0.72.

Normative broadband individual differences.

Neuroticism and aspects.

Neuroticism was not found to be a significant predictor of startle, nor were any Neuroticism x Stimulus interactions significant. The equivalent models that separately modeled the Neuroticism aspects, Withdrawal and Volatility, also did not contain significant effects. Finally, for models in which both aspects were included, there were no significant Aspect x Stimulus or Aspect x Stimulus² interactions.

Multi-trait.

The model containing Neuroticism, Conscientiousness, and Extraversion modeled simultaneously without interaction terms (i.e., main effects model) did not yield any significant Trait effects. For two-way interaction models, there was a significant Extraversion x Stimulus² interaction while including Neuroticism x Stimulus² and Conscientiousness x Stimulus² terms and the associated interactions with the linear Stimulus term, $\beta = 0.24$, 95% CI [0.02 – 0.45], $t(9720.91) = 2.10$, $p = 0.035$. As can be seen in Figure 7, lower Extraversion is associated with an overall shallower decline from CS+, which appears to be driven by elevated responding to the GS2 to the oCS-.

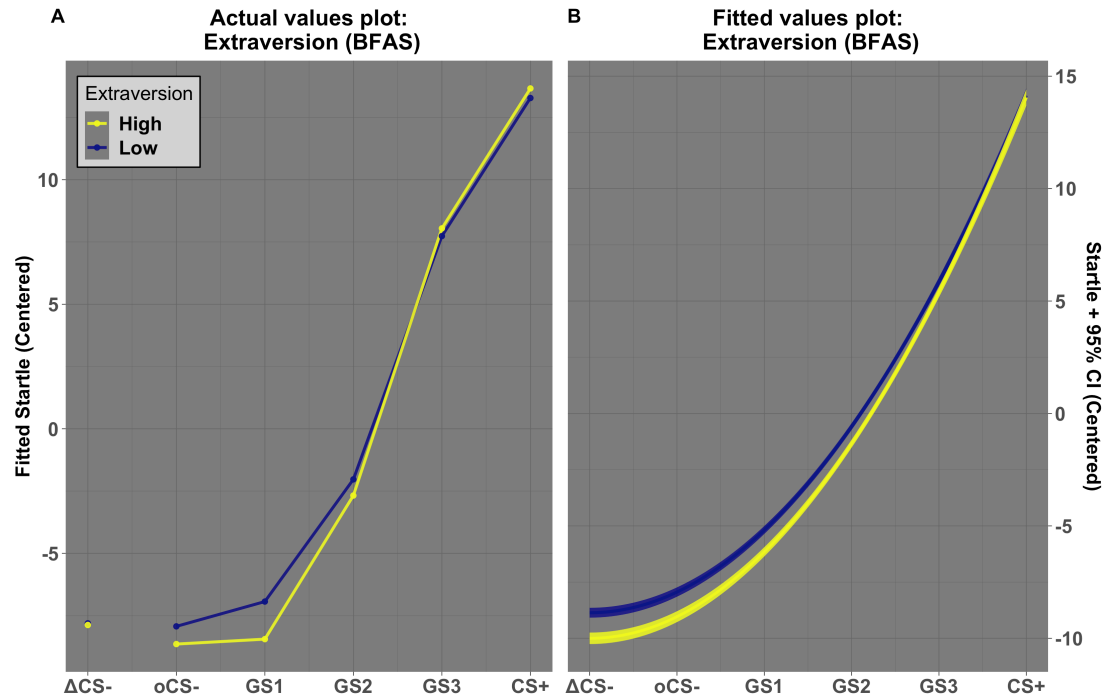


Figure 6. Visual representation of the Extraversion x Stimulus² (quadratic) interaction, with both actual startle EMG values (A) and fitted startle EMG values (B) represented in separate plots. For graphing purposes, 1) a median split was used to plot mean values for those with high or low Extraversion scores, and 2) startle values have been centered within cluster (CWC) so that all participants have a mean of zero. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine (visible as semi-transparent borders around each slope). BFAS = Big Five Aspect Scale; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus.

Multi-aspect.

The model containing Withdrawal and Volatility (Neuroticism aspects), Industriousness and Orderliness (Conscientiousness aspects), and Enthusiasm and Assertiveness (Extraversion aspects) modeled simultaneously without interaction terms (i.e., main effects model) did not yield any significant Aspect effects, nor were there any significant two-way (Aspect x Stimulus dimension) interactions.

Pathological broadband individual differences.

Negative Affectivity.

Negative Affectivity was not found to be a significant predictor of startle, nor were any Negative Affectivity x Stimulus interactions significant.

Multi-trait.

The model containing Negative Affectivity, Disinhibition, and Detachment modeled simultaneously without interaction terms (i.e., main effects model) did not yield any significant Trait effects, nor were there any significant Trait x Stimulus interactions.

Combined PID-5/BFAS “full-spectrum” broadband individual differences (PB).

Distress-PB.

Distress-PB was not found to be a significant predictor of startle, nor were any Distress-PB x Stimulus interactions significant.

Multi-trait.

The model containing Distress-PB, Industriousness-PB, Orderliness-PB, Assertiveness-PB, and Enthusiasm-PB modeled simultaneously without interaction terms (i.e., main effects model) did not yield any significant Trait main effects. In the model in which linear and quadratic interaction terms were modeled together, there was a significant Assertiveness-PB x Stimulus² interaction, $\beta = 0.26$, 95% CI [0.05 – 0.48], $t(9721.41) = 2.37$, $p = 0.0181$, which can be seen in Figure 8. This interaction appears similar to the documented Extraversion x Stimulus² interaction in terms of greater responding to the GS2 through oCS- in those lower on Assertiveness-PB, which is expected given that the Assertiveness-PB variable is partially comprised of Extraversion variance. However, it appears that this interaction has resulted in an even shallower slope for those lower on Assertiveness-PB, as there appears to be a flatter response slope from

the CS+ to the GS3 associated with lower Assertiveness-PB. All other interactions in this model and all interactions in the linear only model were not significant.

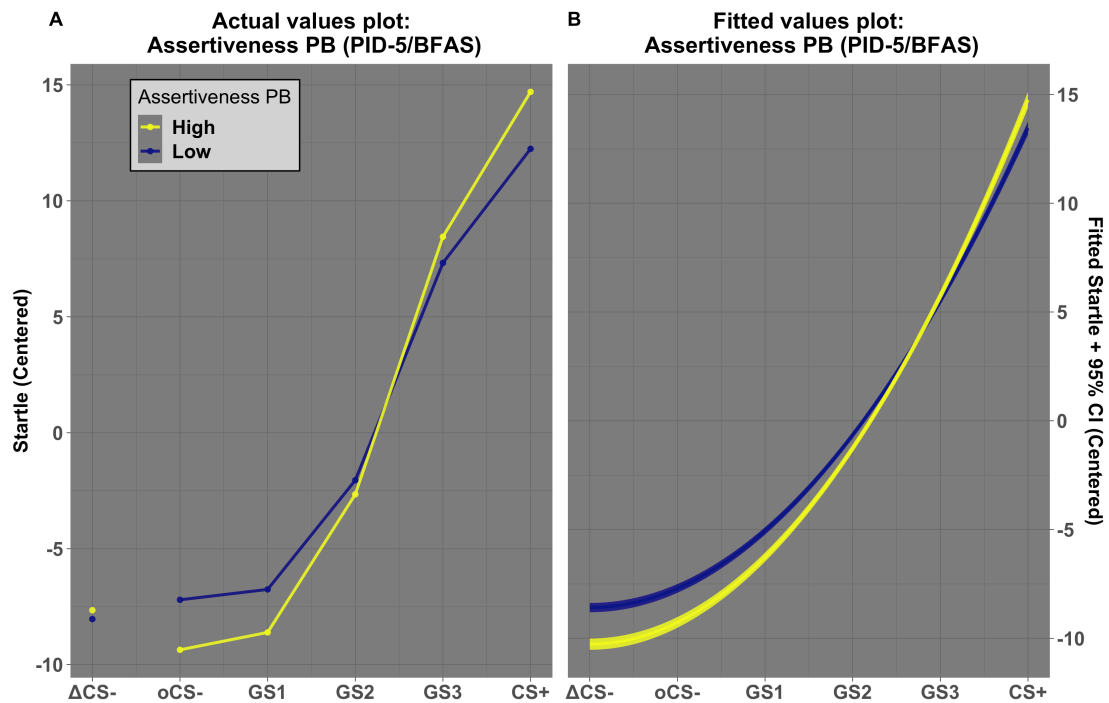


Figure 7. Visual representation of the Assertiveness-PB x Stimulus² (quadratic) interaction, with both actual startle EMG values (A) and fitted startle EMG values (B) represented in separate plots. For graphing purposes, 1) a median split was used to plot mean values for those with high or low Assertiveness-PB scores, and 2) startle values have been centered within cluster (CWC) so that all participants have a mean of zero. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine (visible as semi-transparent borders around each slope). BFAS = Big Five Aspect Scale; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; PB = PID-5/BFAS Composite; PID-5 = Personality Inventory for DSM-5.

Narrowband individual differences.

None of the main effect models for any of the tested narrowband individual differences (TF-44, ASI, IUSF, STAI-T) yielded significant Trait predictors. There was a significant TF-44 x Stimulus², $\beta = -0.01$, 95% CI [0.02 – 0.00], $t(9721.19) = -2.07$, $p = 0.039$, and a significant IUSF x Stimulus² interaction, $\beta = -0.30$, 95%, CI [-0.50 – 0.11], $t(9721.35) = -3.01$, $p = 0.003$. Both interactions capture similar effects, as can be seen in

Figure 9 and Figure 10: there was a shallower decline from CS+ to GS1 in those higher on TF-44I/USF. The linear interactions for TF-44 and IUSF were not significant, as were all interaction terms for ASI and STAI-T.

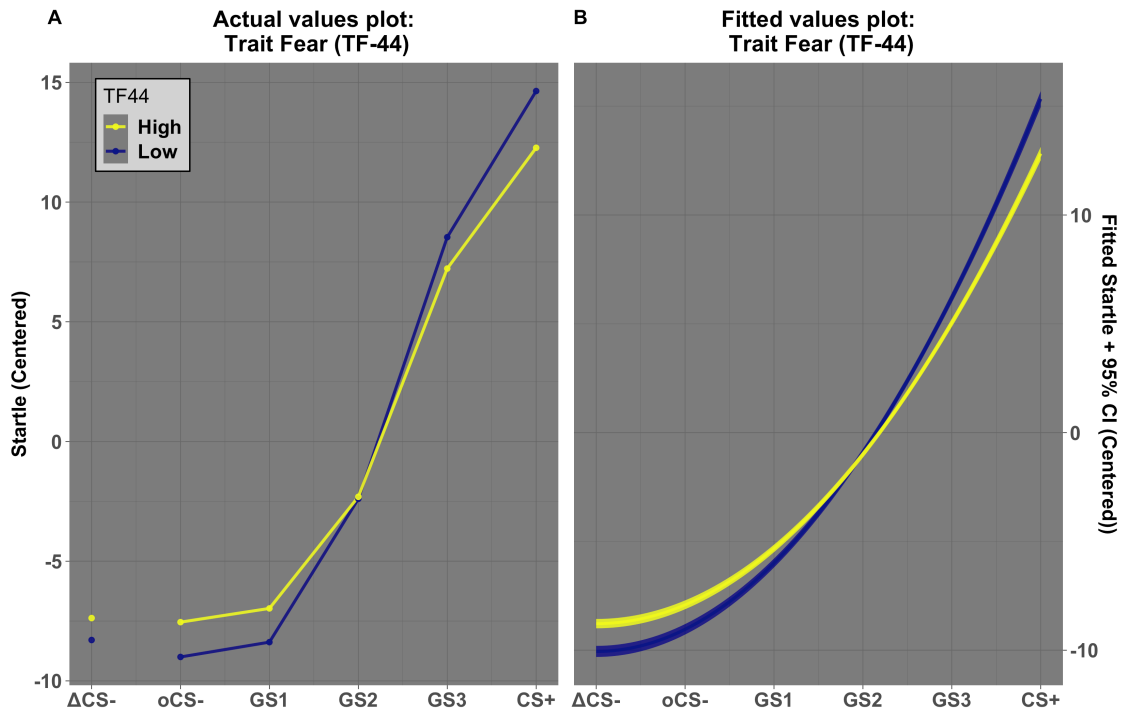


Figure 8. Visual representation of the TF-44 x Stimulus² (quadratic) interaction, with both actual startle EMG values (A) and fitted startle EMG values (B) represented in separate plots. For graphing purposes, 1) a median split was used to plot mean values for those with high or low TF-44 scores, and 2) startle values have been centered within cluster (CWC) so that all participants have a mean of zero. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine (visible as semi-transparent borders around each slope). CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; TF-44 = Trait Fear 44-item.

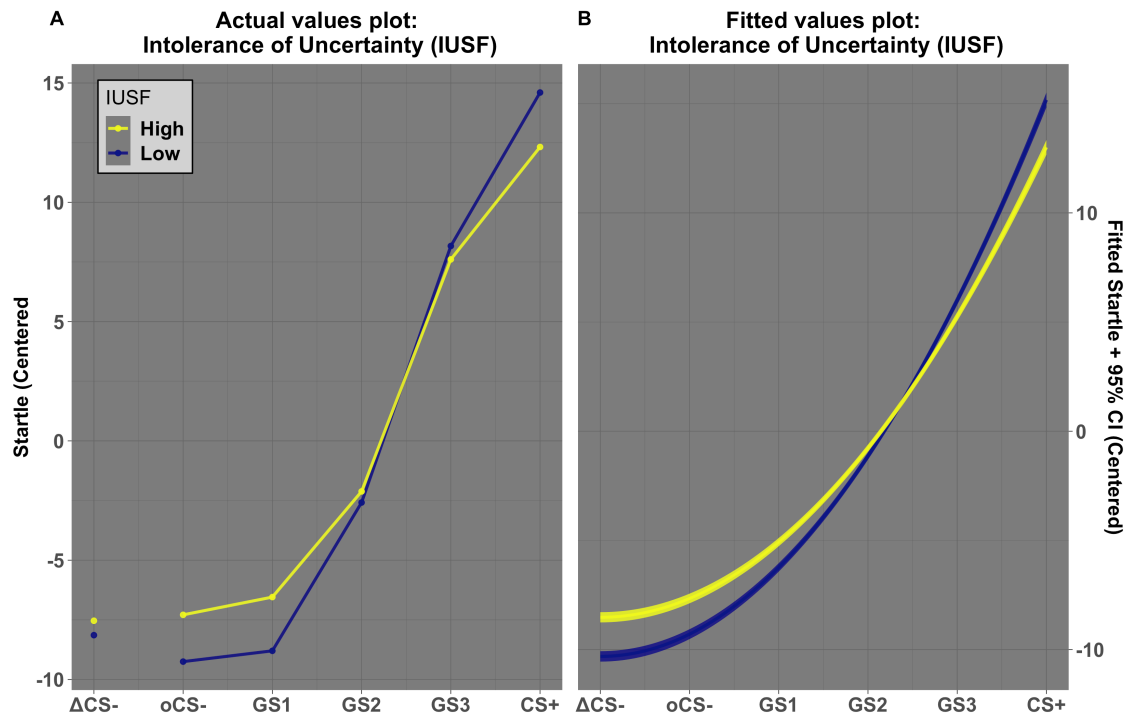


Figure 9. Visual representation of the IUSF x Stimulus² (quadratic) interaction, with both actual startle EMG values (A) and fitted startle EMG values (B) represented in separate plots. For graphing purposes, 1) a median split was used to plot mean values for those with high or low IUSF scores, and 2) startle values have been centered within cluster (CWC) so that all participants have a mean of zero. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine (visible as semi-transparent borders around each slope). CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; IUSF = Intolerance of Uncertainty – Short Form.

A1: Risk ratings.

Testing of the null-model indicated a multilevel model is borderline appropriate for these data ($ICC = .05$). Initial model fitting resulted in Model 2 (random-intercept, random-slope for Stimulus dimension, fixed-effect for Stimulus² dimension) as the most acceptable model for the following risk rating analyses based on our established criteria, with both the random-intercept and all Stimulus effects contributing significantly to the model and adjusted $ICC = .21$.

Normative broadband individual differences.

Neuroticism and aspects.

Neuroticism was a significant predictor of risk ratings, $\beta = 0.025$, 95% CI [0.001 – 0.049], $t(285.56) = 2.08$, $p = 0.038$, with higher Neuroticism associated with higher risk ratings. There was a significant Neuroticism x Stimulus interaction, $\beta = 0.012$, 95% CI [0.0001 – 0.0239], $t(286.46) = 1.979$, $p = 0.049$. As can be seen in Figure 11, those higher on Neuroticism have a slightly shallower decrease from the CS+ to the GS3, indicating a moderate degree of generalization associated with higher Neuroticism. The interaction with the quadratic Stimulus term was not significant. The equivalent models that separately modeled the Neuroticism aspects (Withdrawal and Volatility) did not contain significant main effects or interactions, nor did models that modeled the Aspects simultaneously.

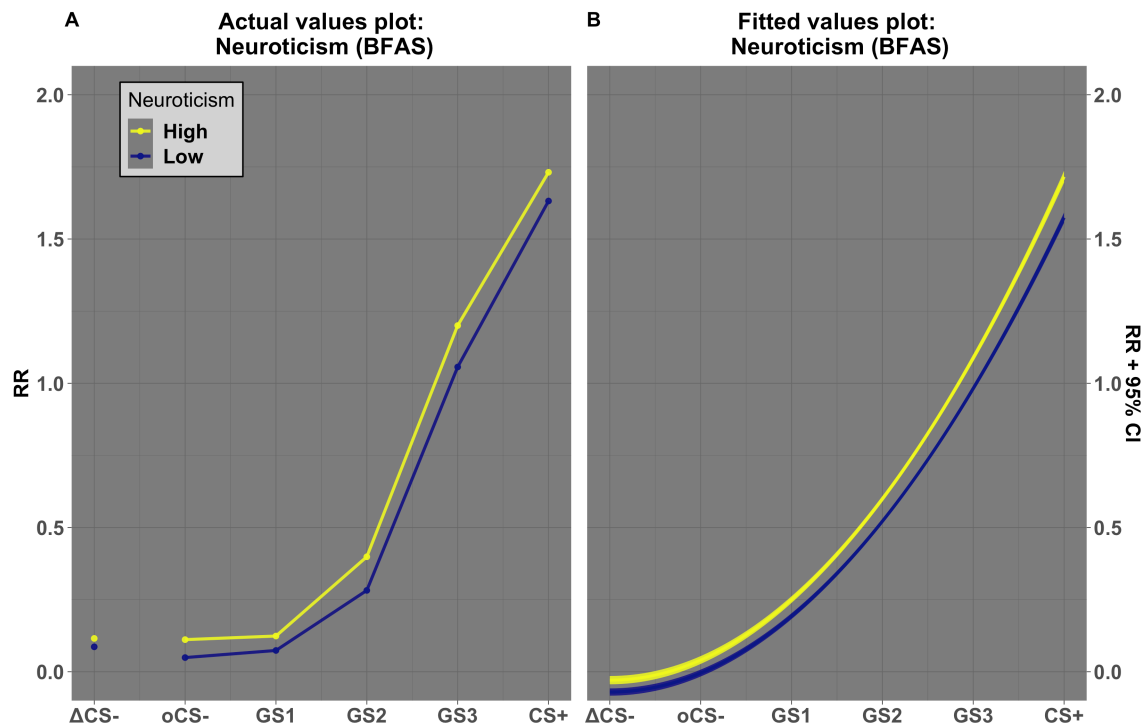


Figure 10. Visual representation of the Neuroticism x Stimulus interaction, with both actual risk rating values (A) and fitted risk rating values (B) represented in separate plots.

For graphing purposes, a median split was used to plot mean values for those with high or low Neuroticism scores. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine 95% confidence intervals (visible as semi-transparent borders around each slope). CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; IUSF = Intolerance of Uncertainty – Short Form; RR = risk ratings.

Multi-trait.

The model containing Neuroticism, Conscientiousness, and Extraversion modeled simultaneously without interaction terms (i.e., main effects model) did not yield any significant Trait main effects. For two-way interaction models, there was a significant Neuroticism x Stimulus interaction while including the Extraversion x Stimulus and Conscientiousness x Stimulus terms in the model, $\beta = 0.163$, 95% CI [0.0035 – 0.0290], $t(9720.91) = 2.5$, $p = 0.0013$. This interaction was visually and statistically close to identical to the previously documented Neuroticism x Stimulus interaction in the model without the other personality variables, and therefore is not interpreted further in this section. All other interactions were not significant.

Multi-aspect.

The model containing Withdrawal and Volatility (Neuroticism aspects), Industriousness and Orderliness (Conscientiousness aspects), and Enthusiasm and Assertiveness (Extraversion aspects) modeled simultaneously without interaction terms (i.e., main effects model) did not yield any significant Aspect effects, nor were there any significant two-way (Aspect x Stimulus dimension) interactions.

Pathological broadband individual differences.

Negative Affectivity.

Negative Affectivity was not found to be a significant predictor of risk ratings, nor were any Negative Affectivity x Stimulus interactions significant.

Multi-trait.

The model containing Negative Affectivity, Disinhibition, and Detachment modeled simultaneously without interaction terms (i.e., main effects model) yielded a significant Negative Affectivity main effect, $\beta = 0.0442$, 95% CI [0.0019 – 0.0866], $t(285.14) = 2.04$, $p = 0.042$, with increases in Negative Affectivity associated with elevated risk ratings. There were no main effects of Disinhibition or Detachment.

The Negative Affectivity x Stimulus interaction was significant, $\beta = -0.021$, 95% CI [0.0048 – 0.0469], $t(286.56) = 2.41$, $p = 0.017$. As can be seen in Figure 12, those higher on Negative Affectivity have a shallower decrease in responding from the CS+ to the GS3. The Detachment x Stimulus interaction was also significant, $\beta = -0.0210$, 95% CI [-0.0394 – -0.0026], $t(286.27) = -2.2337$, $p = 0.026$, which in Figure 13 is seen at the GS2 and GS1: those higher on Detachment were slightly elevated on the GS2 relative to those lower on Detachment, and this is reversed for the GS1, resulting in a slightly shallower decline from GS2 to GS1 for those higher on Detachment. However, the CIs in the fitted values plot overlap considerably, indicating that the shallowed decline seen in the actual values plot is potentially not evident in the fitted values. The Disinhibition x Stimulus and all interactions with the quadratic Stimulus term were not significant.

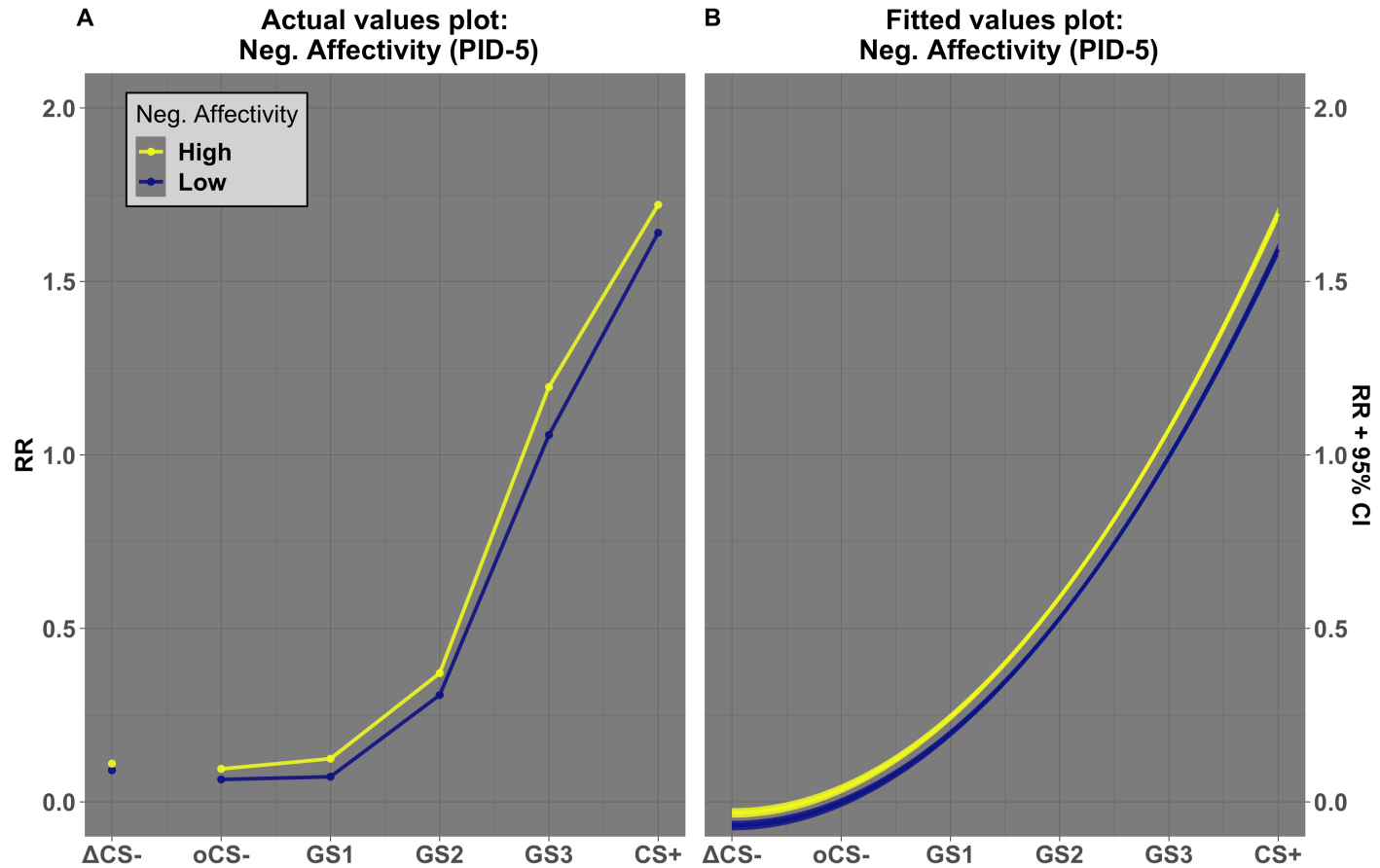


Figure 11. Visual representation of the Negative Affectivity x Stimulus interaction, while controlling for Detachment and Disinhibition, with both actual risk rating values (A) and fitted risk rating values (B) represented in separate plots. For graphing purposes, a median split was used to plot mean values for those with high or low Negative Affectivity scores. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine 95% confidence intervals (visible as semi-transparent borders around each slope). CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; PID-5 = Personality Inventory for DSM-5; RR = risk ratings.

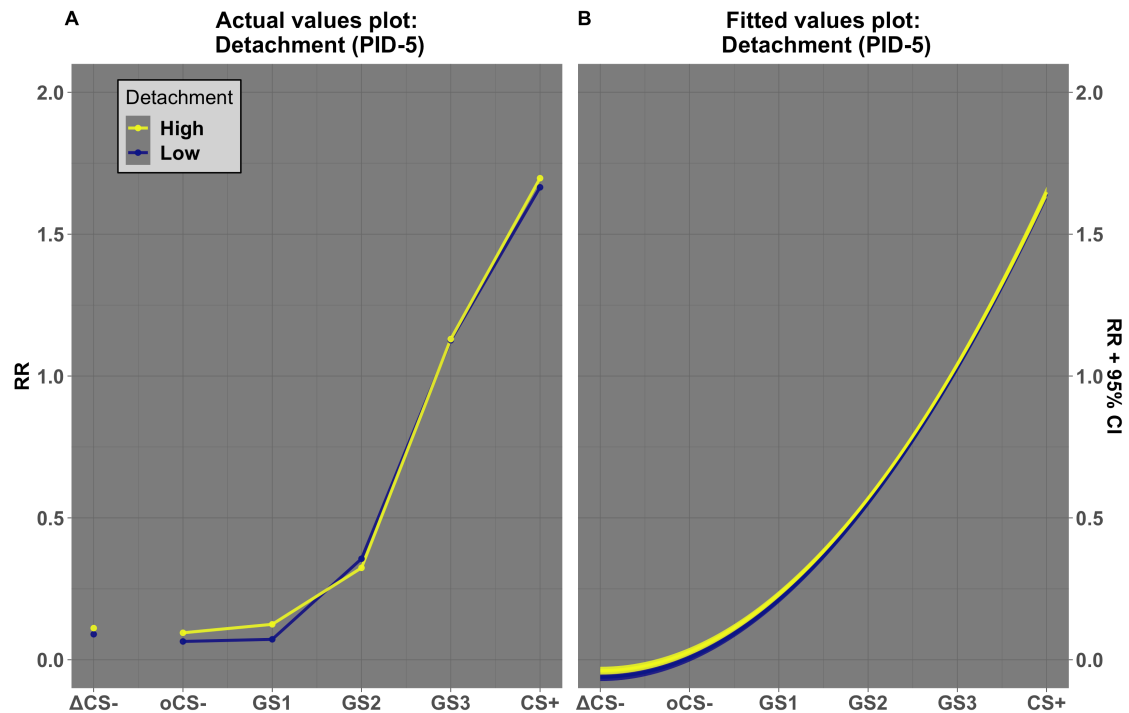


Figure 12. Visual representation of the Detachment x Stimulus interaction while controlling for Negative Affectivity and Disinhibition, with both actual risk rating values (A) and fitted risk rating values (B) represented in separate plots. For graphing purposes, a median split was used to plot mean values for those with high or low Detachment scores. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine 95% confidence intervals (visible as semi-transparent borders around each slope). CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; PID-5 = Personality Inventory for DSM-5; RR = risk ratings.

Combined PID-5/BFAS “full-spectrum” broadband individual differences

(PB).

Distress-PB.

Distress-PB was not found to be a significant predictor of risk ratings, nor were any Distress-PB x Stimulus interactions significant.

Multi-trait.

The model containing Distress-PB, Industriousness-PB, Orderliness-PB, Assertiveness-PB, and Enthusiasm-PB modeled simultaneously without interaction terms (i.e., main effects model) yielded a significant Distress-PB main effect, $\beta = 0.0388$, 95%

CI [0.0069 – 0.0707], $t(285.16) = 2.3809$, $p = 0.018$, with increased Distress-PB associated with elevated risk ratings. For two-way interaction models, there was a significant Distress-PB x Stimulus interaction, $\beta = 0.0213$, 95% CI [0.0057 – 0.0368], $t(286.96) = 2.6824$, $p = 0.008$, which can be seen in Figure 14 as a subtly shallower decline from CS+ to GS3 in those higher on Distress-PB that resembles the previously documented Negative Affectivity and Neuroticism interactions. There was also a significant Enthusiasm-PB x Stimulus interaction, $\beta = 0.0153$, 95% CI [0.0016 – 0.0291], $t(286.63) = 2.37$, $p = 0.0029$, which resembled the previously documented Detachment interaction, with those lower on Enthusiasm-PB (i.e., those who are likely higher on Detachment) showing a shallower decline from GS2 to GS1 (see Figure 15). Also in line with the previously documented Detachment interaction, the CIs in the fitted values plot for Enthusiasm-PB overlap considerably, indicating that the shallowed decline seen in the actual values plot is potentially not evident in the fitted values. All other interactions were not significant.

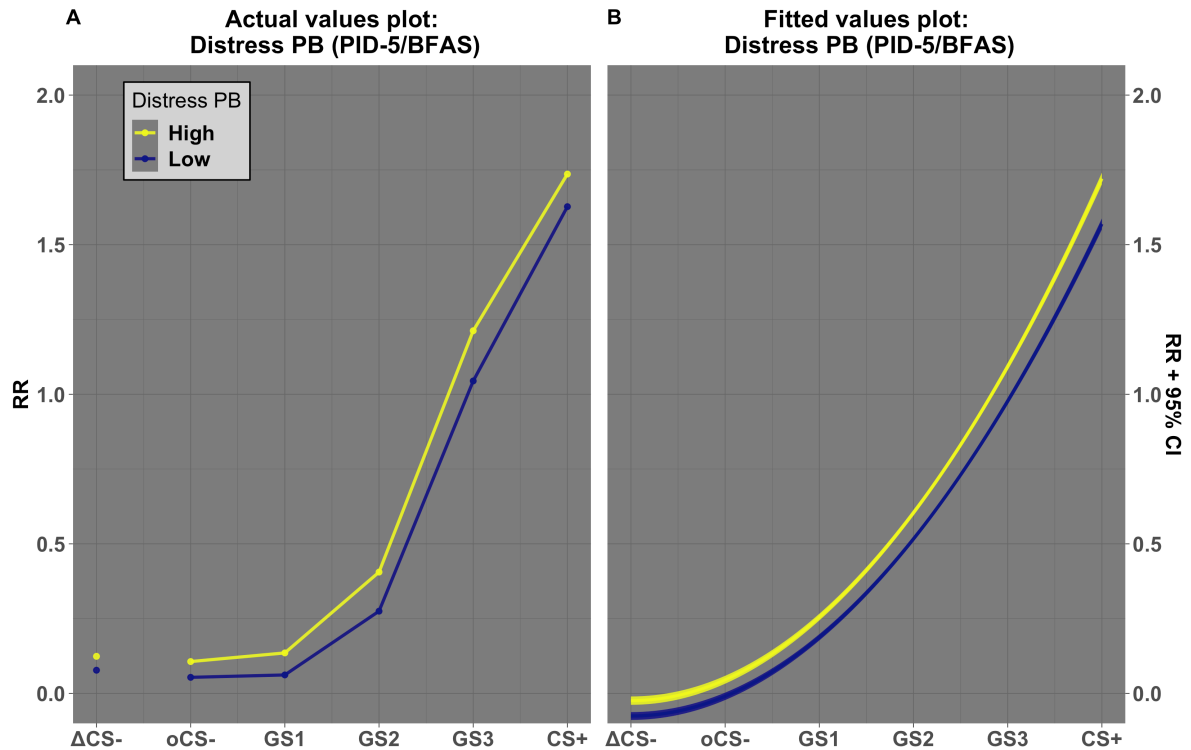


Figure 13. Visual representation of the Distress-PB x Stimulus interaction while controlling for all other “PB” composite variables, with both actual risk rating values (A) and fitted risk rating values (B) represented in separate plots. For graphing purposes, a median split was used to plot mean values for those with high or low Distress-PB scores. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine 95% confidence intervals (visible as semi-transparent borders around each slope). BFAS = Big Five Aspect Scale; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; PB: PID-5/BFAS composite; PID-5 = Personality Inventory for DSM-5; RR = risk ratings.

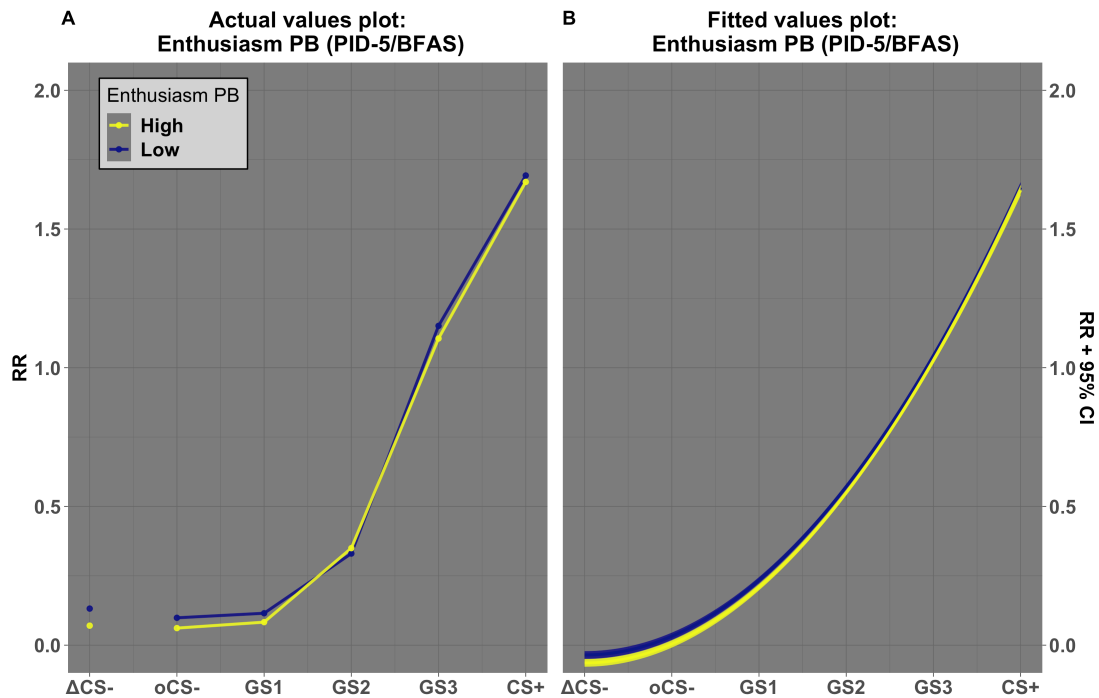


Figure 14. Visual representation of the Enthusiasm-PB x Stimulus interaction while controlling for all other “PB” composite variables, with both actual risk rating values (A) and fitted risk rating values (B) represented in separate plots. For graphing purposes, a median split was used to plot mean values for those with high or low Enthusiasm-PB scores. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine 95% confidence intervals (visible as semi-transparent borders around each slope). BFAS = Big Five Aspect Scale; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; PB: PID-5/BFAS composite; PID-5 = Personality Inventory for DSM-5; RR = risk ratings.

Narrowband individual differences.

There was a significant main effect of TF-44, $\beta = 0.0012$, 95% CI

[0.0001 – 0.0023], $t(286.99) = 2.0945$, $p = 0.0037$, and IUSF, $\beta = 0.0034$, 95% CI

[0.0008 – 0.0061], $t(286.54) = 2.5634$, $p = 0.0011$, on risk ratings; for both effects the

narrowband trait was positively related to risk ratings such that increased levels of the

trait were associated with elevated risk ratings. The ASI and STAI-T main effects were

not significant. There was a significant IUSF x Stimulus² interaction, $\beta = 0.0039$, 95% CI

[0.0005 – 0.0072], $t(4103.47) = 2.2783$, $p = 0.0023$, with those higher on IUSF showing a

slightly shallower decline in responding from CS+ to GS3 (see Figure 16). All other narrowband interactions were not significant.

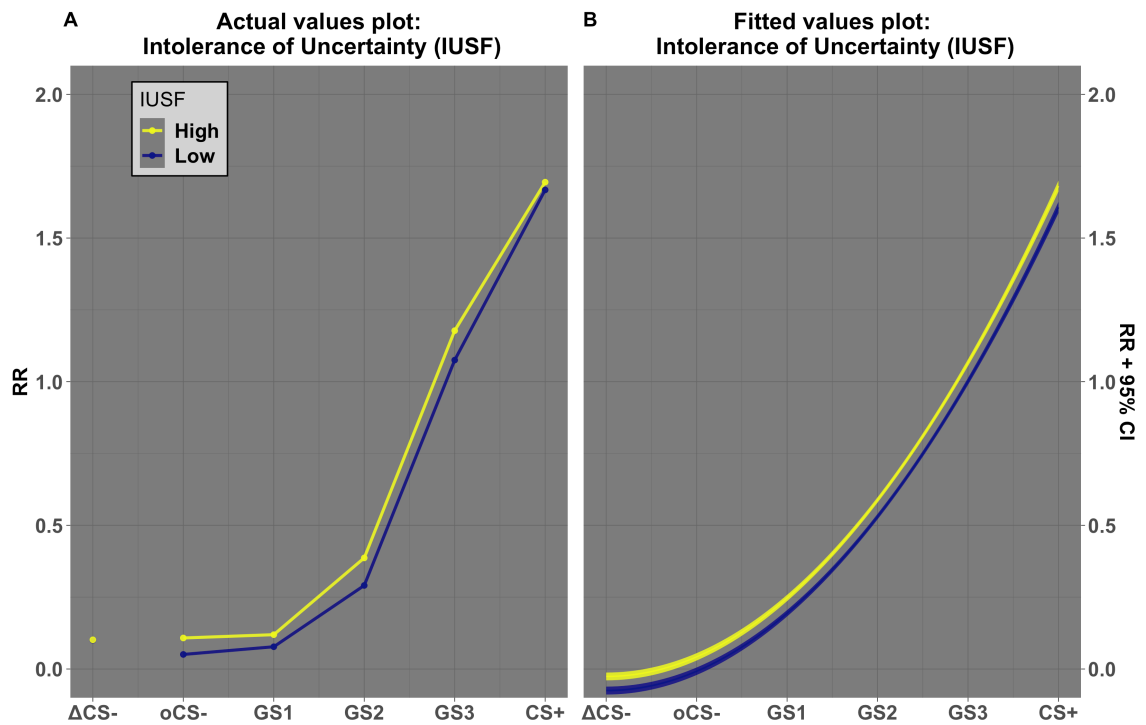


Figure 15. Visual representation of the IUSF x Stimulus² interaction, with both actual risk rating values (A) and fitted risk rating values (B) represented in separate plots. For graphing purposes, a median split was used to plot mean values for those with high or low IUSF scores. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine 95% confidence intervals (visible as semi-transparent borders around each slope). CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; IUSF = Intolerance of Uncertainty – Short Form; RR = risk ratings.

A1: Avoidance.

Normative broadband individual differences.

Neuroticism and aspects.

Neuroticism alone was not a significant predictor of avoidance, however, there was a significant Neuroticism x Stimulus² interaction, $\beta = 1.0466$, 95% CI [1.0054 – 1.0895], $\chi^2(1, N=10300) = 4.937$, $p = 0.026$. As seen in Figure 17, this appears to reflect a non-generalization effect, with those higher on Neuroticism avoiding the CS+

and GS3 at a higher rate than those lower on Neuroticism, but without a corresponding change in the grade of the decline from CS+ to GS3. All other interactions were not significant.

The equivalent models that separately or simultaneously modeled the Neuroticism aspects (Withdrawal and Volatility) did not contain significant main effects. However, as can be seen in Figure 18, there was a significant Volatility x Stimulus² interaction in the model that only included Volatility, $\beta = 1.0481$, 95% CI [1.0075 – 1.08903], $\chi^2(1, N=10300) = 5.4265$, $p = 0.020$. All other Volatility interactions, as well as Withdrawal interactions, were not significant.

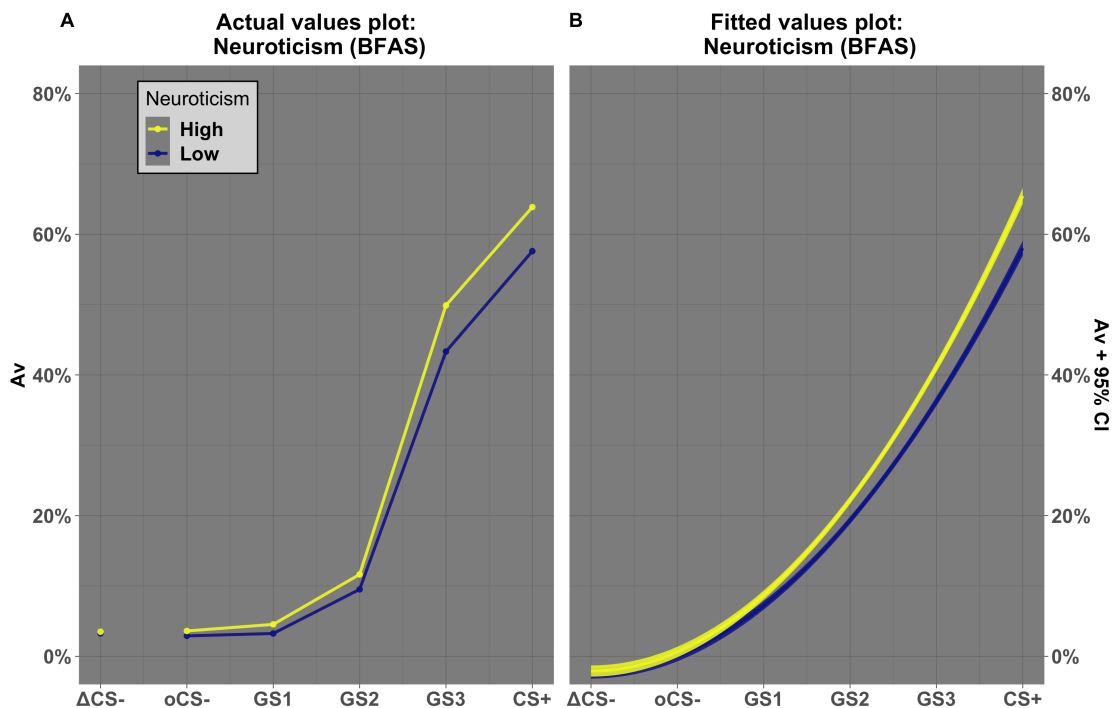


Figure 16. Visual representation of the Neuroticism x Stimulus² interaction, with both actual avoidance % values (A) and fitted avoidance % (B) represented in separate plots. For graphing purposes, a median split was used to plot mean values for those with high or low Neuroticism scores. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine 95% confidence intervals (visible as semi-transparent borders around each slope). Av = avoidance; BFAS = Big Five Aspect Scale; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus.

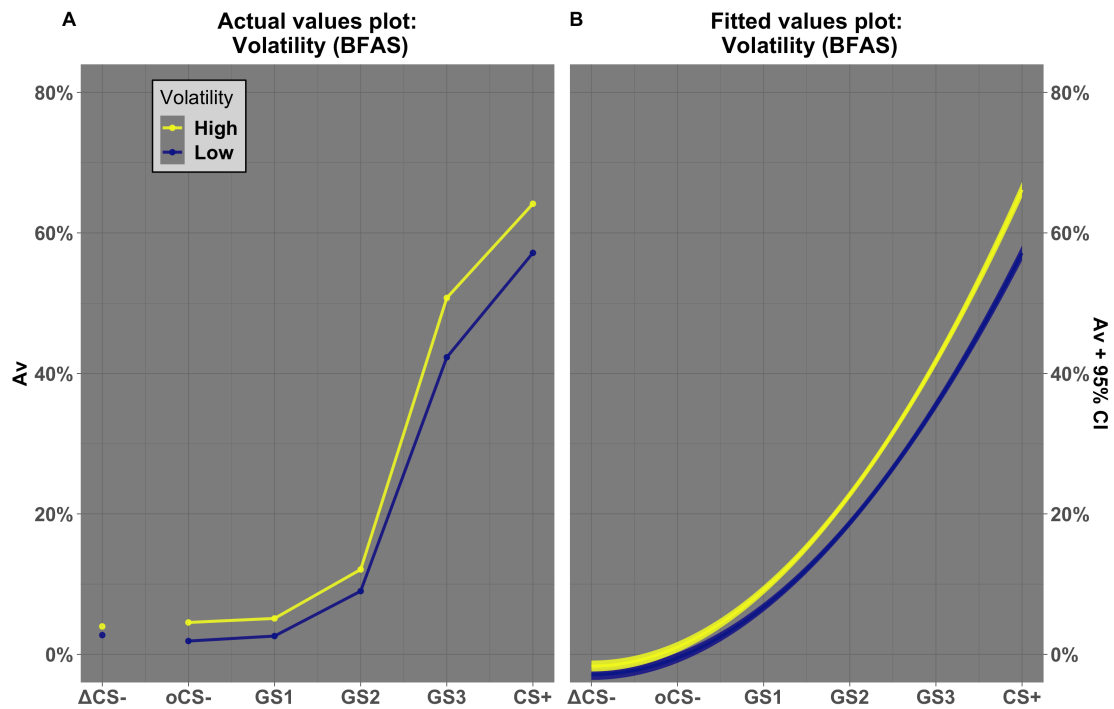


Figure 17. Visual representation of the Volatility x Stimulus² interaction, with both actual avoidance % values (A) and fitted avoidance % (B) represented in separate plots. For graphing purposes, a median split was used to plot mean values for those with high or low Volatility scores. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine 95% confidence intervals (visible as semi-transparent borders around each slope). Av = avoidance; BFAS = Big Five Aspect Scale; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus.

Conscientiousness and aspects.

Conscientiousness alone was not a significant predictor of avoidance, however, there was a significant Conscientiousness x Stimulus² interaction, $\beta = 0.9557$, 95% CI [0.9176 – 0.9954], $\chi^2(1, N=10300) = 4.7606$, $p = 0.029$. This appears to be driven by those who are higher on Conscientiousness showing a shallower decline from CS+ to GS3 (see Figure 19). The Conscientiousness interaction with the linear Stimulus term was not significant. The equivalent models that separately or simultaneously modeled the Conscientiousness aspects (Industriousness and Orderliness) did not contain significant

main effects. However, there was a significant Orderliness x Stimulus² interaction in the model that only included Orderliness, $\beta = 0.9539$, 95% CI [0.9168 – 0.9924], $\chi^2(1, N=10300) = 5.4606$, $p = 0.019$, which resembled the previously documented Conscientiousness x Stimulus² interaction, as those higher on Orderliness showed a shallower decline from CS+ to GS3; also notable is that those with higher Orderliness had overall higher avoidance rates at this segment of the response slope, whereas when looking at the trait level, lower Conscientiousness was related to overall higher avoidance rates for this segment (see Figure 20). All other Orderliness interactions, as well as Industriousness interactions, were not significant.

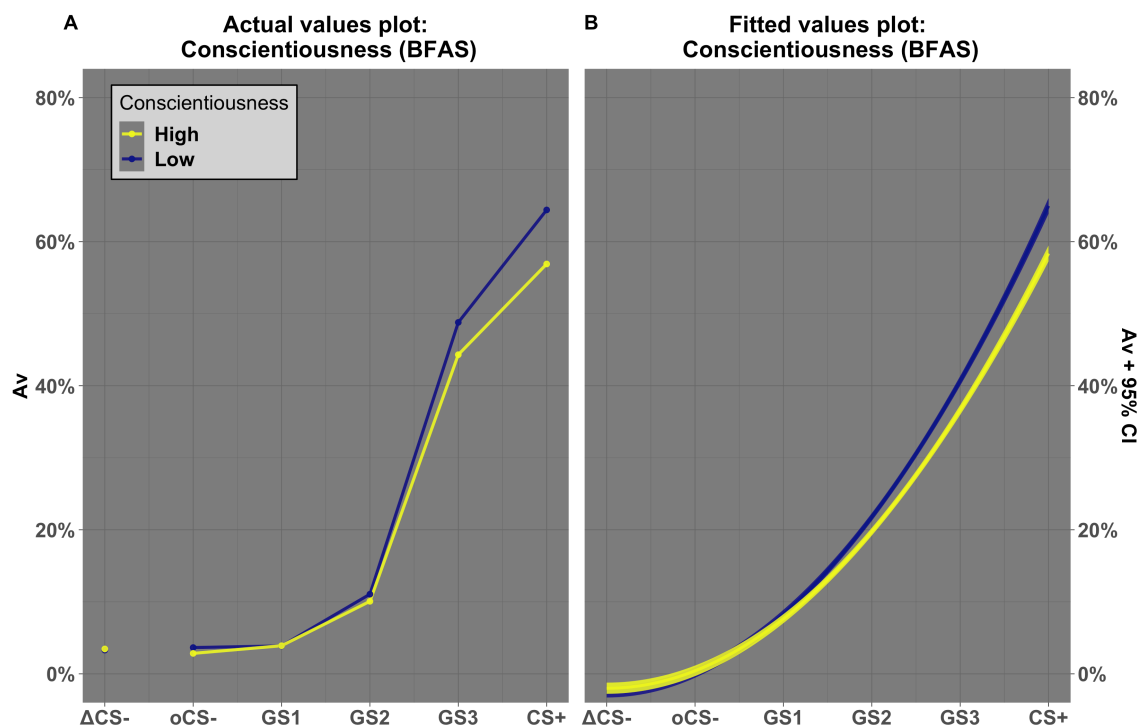


Figure 18. Visual representation of the Conscientiousness x Stimulus² interaction, with both actual avoidance % values (A) and fitted avoidance % (B) represented in separate plots. For graphing purposes, a median split was used to plot mean values for those with high or low Conscientiousness scores. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine 95% confidence intervals (visible as semi-transparent borders around each slope). Av = avoidance; BFAS = Big Five Aspect Scale; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus.

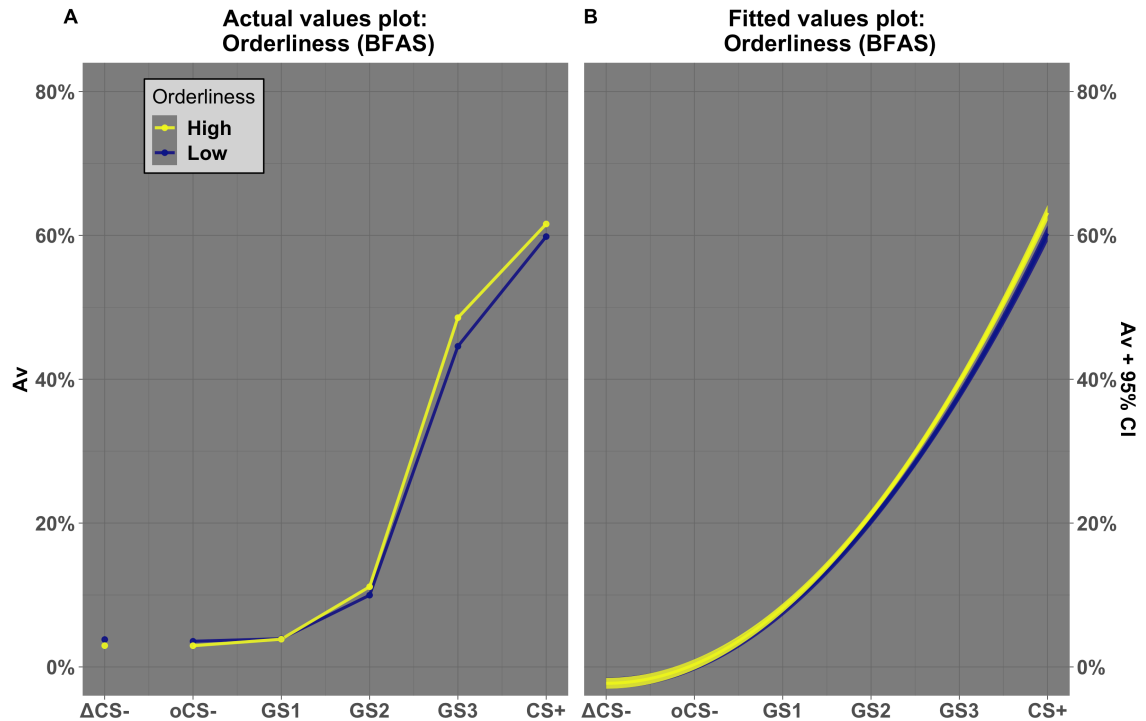


Figure 19. Visual representation of the Orderliness x Stimulus² interaction, with both actual avoidance % values (A) and fitted avoidance % (B) represented in separate plots. For graphing purposes, a median split was used to plot mean values for those with high or low Orderliness scores. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine 95% confidence intervals (visible as semi-transparent borders around each slope). Av = avoidance; BFAS = Big Five Aspect Scale; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus.

Extraversion and aspects.

Extraversion or its aspects (Assertiveness, Enthusiasm) were not found to be a significant predictor of avoidance, nor were any interactions significant.

Multi-trait.

The model containing Neuroticism, Conscientiousness, and Extraversion modeled simultaneously without interaction terms (i.e., main effects model) did not yield any significant Trait effects. For two-way interaction models, there was a significant Conscientiousness x Stimulus² interaction while including the Conscientiousness x

Stimulus interaction and both quadratic and linear interactions terms for Extraversion and Neuroticism in the model, $\beta = 0.9539$, 95% CI [0.9112 – 0.9986], $\chi^2(1, N=10300) = 4.0779$, $p = 0.0013$. This interaction was visually and statistically close to identical to the previously documented Conscientiousness x Stimulus² interaction in the model without the other personality variables, and therefore is not interpreted further in this section. The separate model that tested linear interactions only did not yield any significant interactions.

Multi-aspect.

The model containing Withdrawal and Volatility (Neuroticism aspects), Industriousness and Orderliness (Conscientiousness aspects), and Enthusiasm and Assertiveness (Extraversion aspects) modeled simultaneously without interaction terms (i.e., main effects model) yielded two significant Aspect main effects. In this model, Volatility significantly predicted increased avoidance, $\beta = 1.4014$, 95% CI [1.0446 – 1.8801], $\chi^2(1, N=10300) = 5.0683$, $p = 0.0024$, whereas Assertiveness significantly predicted decreased avoidance, $\beta = 0.7325$, 95% CI [0.5655 – 0.9487], $\chi^2(1, N=10300) = 5.5629$, $p = 0.0018$. There were no significant two-way (Aspect x Stimulus dimension) interactions.

Pathological broadband individual differences.

Negative Affectivity.

Negative Affectivity alone was not found to be a significant predictor of avoidance. Both the Negative Affectivity x Stimulus, $\beta = 0.7668$, 95% CI [0.5992 – 0.9814], $\chi^2(1, N=10300) = 4.4483$, $p = 0.0035$, and Negative Affectivity x Stimulus², $\beta = 1.0436$, 95% CI [1.0050 – 1.0836], $\chi^2(1, N=10300) = 4.9346$, $p = 0.0026$

were significant in the same model used to detect quadratic interactions. As seen in Figure 21, those higher on Negative Affectivity demonstrated a shallower decline in avoidance from CS+ to GS3, as well as from GS2 to oCS- (although CIs in the fitted values plot overlapped at this level, potentially indicating the observed pattern in the actual values plot is not reflected in the fitted values). The separate model testing the linear interaction alone was not significant.

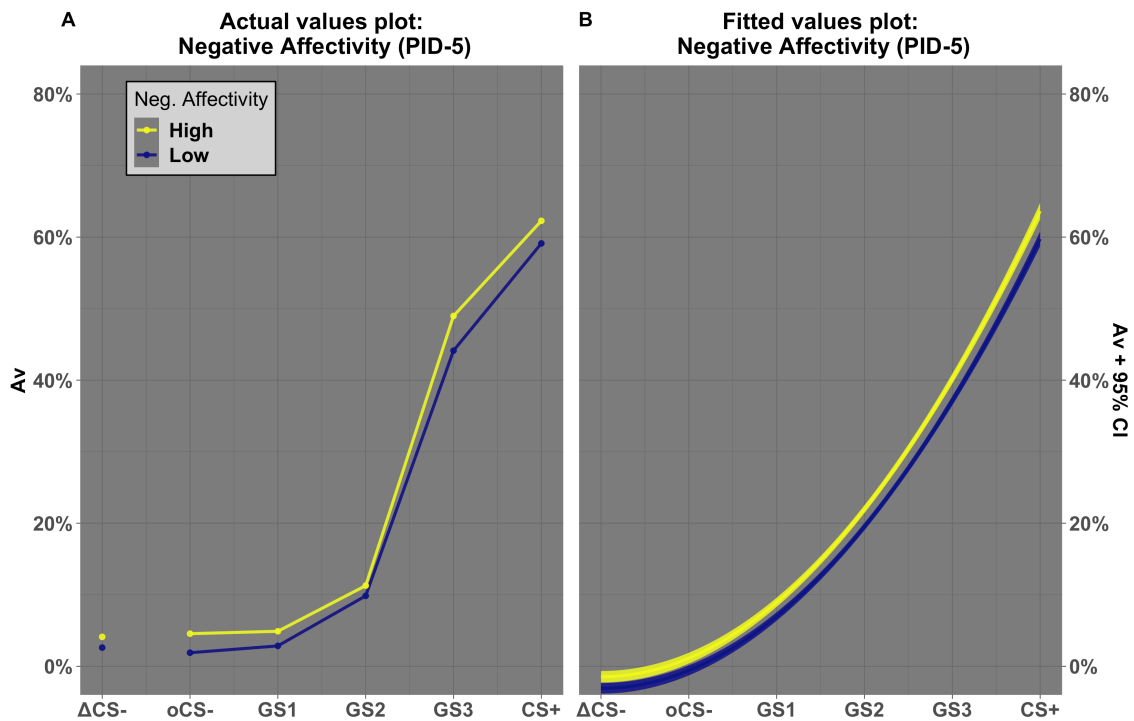


Figure 20. Visual representation of the Negative Affectivity x Stimulus² interaction, with both actual avoidance % values (A) and fitted avoidance % (B) represented in separate plots. For graphing purposes, a median split was used to plot mean values for those with high or low Negative Affectivity scores. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine 95% confidence intervals (visible as semi-transparent borders around each slope). Av = avoidance; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; PID-5 = Personality Inventory for DSM-5

Disinhibition.

Disinhibition alone was not found to be a significant predictor of avoidance, nor were any of the interaction terms significant.

Detachment.

Detachment alone was not found to be a significant predictor of avoidance, nor were any of the interaction terms significant.

Multi-trait.

The model containing Negative Affectivity, Disinhibition, and Detachment modeled simultaneously without interaction terms (i.e., main effects model) yielded a significant Negative Affectivity main effect, $\beta = 1.5500$, 95% CI [1.0615 – 2.2633], $\chi^2(1, N=10300) = 5.1474$, $p = 0.0023$, with increases in Negative Affectivity associated with an increased chance of avoidance. There was also a significant Disinhibition main effect, $\beta = 0.7639$, 95% CI [0.5854 – 0.9968], $\chi^2(1, N=10300) = 3.9350$, $p = 0.0047$, with increases in Disinhibition associated with a decreased chance of avoidance (i.e., increased approach). The Detachment main effect was not significant.

In the model with both linear and quadratic terms, the Negative Affectivity x Stimulus, $\beta = 0.5147$, 95% CI [0.3147 – 0.8418], $\chi^2(1, N=10300) = 7.0007$, $p = 0.008$, and Negative Affectivity x Stimulus² interactions were significant, $\beta = 1.1215$, 95% CI [1.0416 – 1.2075], $\chi^2(1, N=10300) = 9.242$, $p = 0.002$. These interactions were visually and statistically close to identical to the previously documented Negative Affectivity interactions in the model without the other personality variables, and therefore is not interpreted further in this section. Additionally, the Detachment x Stimulus, $\beta = 1.1215$, 95% CI [1.0456 – 2.5007], $\chi^2(1, N=10300) = 7.657$, $p = 0.031$, and Detachment x Stimulus² interactions were significant, $\beta = 0.9127$, 95% CI [0.8555 – 0.9737], $\chi^2(1, N=10300) = 9.242$, $p = 0.006$. As seen in Figure 22, those higher on Detachment show a

slightly shallower decline in responding from GS3 to GS2, as well as a more markedly shallow decline from GS2 to oCS-.

The Disinhibition x Stimulus interactions from this model and all interaction terms from the model which only includes linear interactions were not significant.

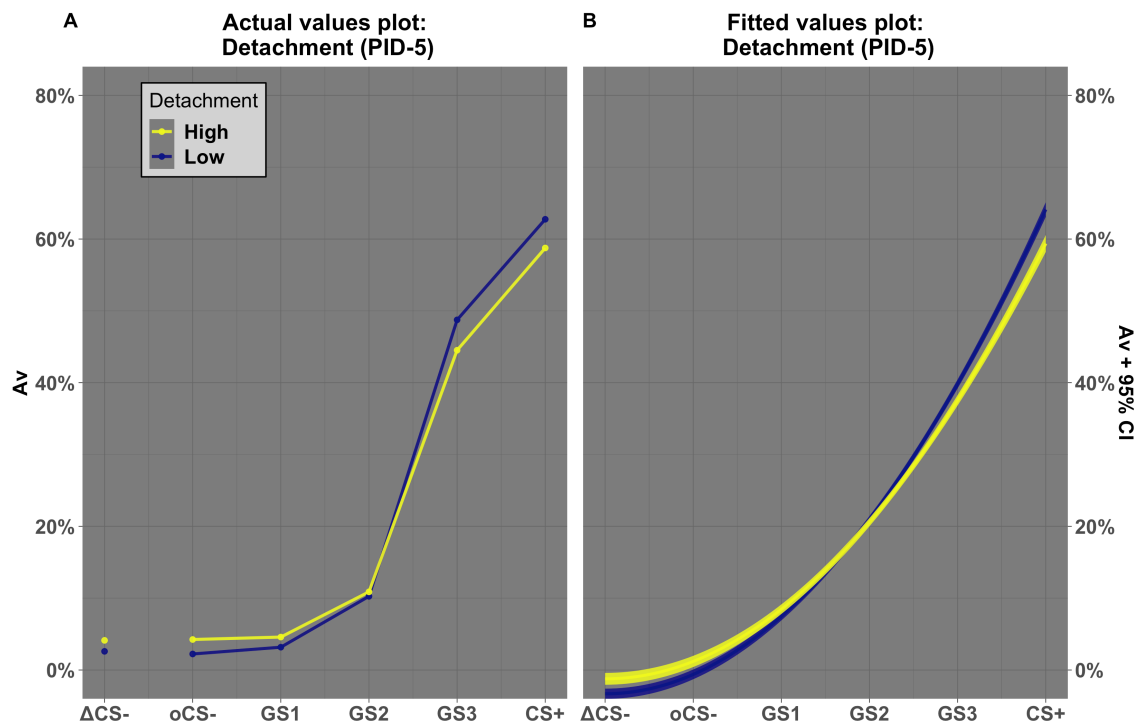


Figure 21. Visual representation of the Detachment x Stimulus² interaction while controlling for Negative Affectivity and Disinhibition, with both actual avoidance % values (A) and fitted avoidance % (B) represented in separate plots. For graphing purposes, a median split was used to plot mean values for those with high or low Detachment scores. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine 95% confidence intervals (visible as semi-transparent borders around each slope). Av = avoidance; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; PID-5 = Personality Inventory for DSM-5

Combined PID-5/BFAS “full-spectrum” broadband individual differences

(PB).

Distress-PB.

Distress-PB alone was not found to be a significant predictor of avoidance. However, both the Distress-PB x Stimulus, $\beta = 0.7359$, 95% CI [0.5740 – 0.9435], $\chi^2(1, N=10300) = 5.9520$, $p = 0.0016$, and Distress-PB x Stimulus², $\beta = 1.0509$, 95% CI [1.0116 – 1.0917], $\chi^2(1, N=10300) = 6.5193$, $p = 0.0011$, interactions were significant in the same model. As seen in Figure 23, those higher on Distress-PB showed a shallower decline in responding from the CS+ to GS3. The separate model testing the linear interaction alone was not significant.

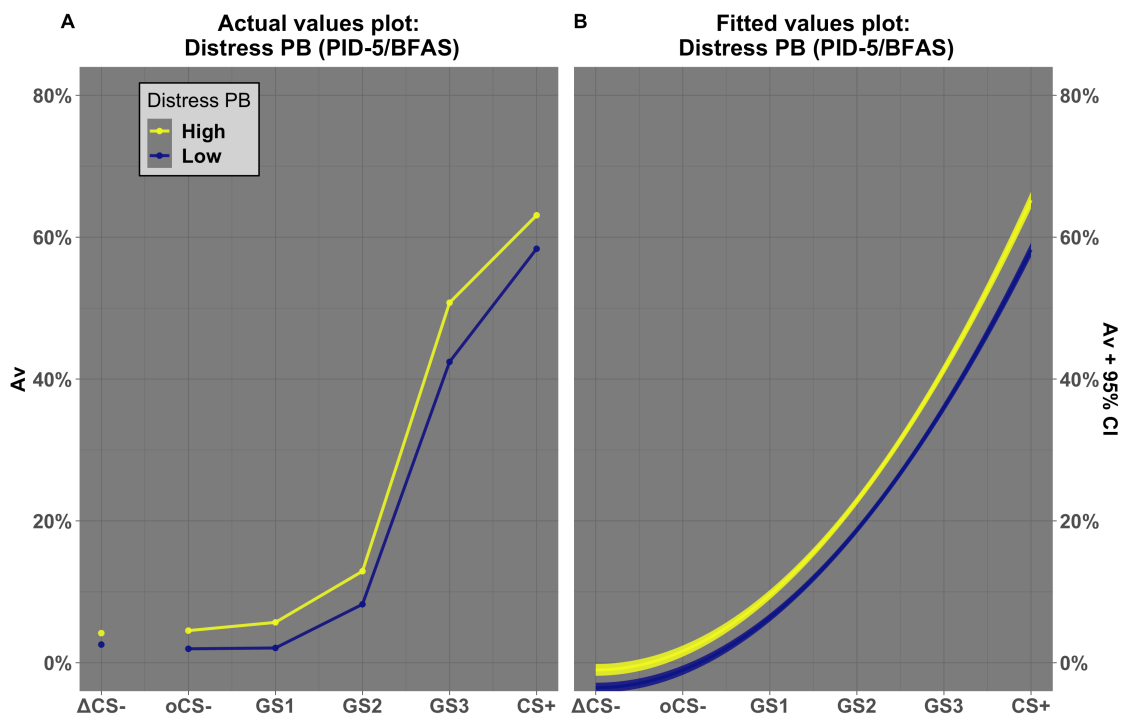


Figure 22. Visual representation of the Distress-PB x Stimulus² interaction, with both actual avoidance % values (A) and fitted avoidance % (B) represented in separate plots. For graphing purposes, a median split was used to plot mean values for those with high or low Distress-PB scores. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine 95% confidence intervals (visible as semi-transparent borders around each slope). Av = avoidance; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; PB = PID-5/BFAS composite; PID-5 = Personality Inventory for DSM-5.

Multi-trait.

The model containing Distress-PB, Industriousness-PB, Orderliness-PB, Assertiveness-PB, and Enthusiasm-PB modeled simultaneously without interaction terms (i.e., main effects model) yielded a significant main effect of Distress-PB, $\beta = 1.4065$, 95% CI [1.0618 – 1.8632], $\chi^2(1, N=10300) = 5.6553$, $p = 0.0017$, with increases in Distress-PB associated with an increased chance of avoidance. There was also a significant Enthusiasm-PB main effect, $\beta = 1.2991$, 95% CI [1.0146 – 1.6634], $\chi^2(1, N=10300) = 4.3035$, $p = 0.0038$, with increases in Enthusiasm-PB also associated with an increased chance of avoidance. All other main effects were not significant.

In the model with both linear and quadratic interaction terms, the Distress-PB x Stimulus, $\beta = 0.6321$, 95% CI [0.4467 – 0.8944], $\chi^2(1, N=10300) = 6.7070$, $p = 0.010$, and Distress-PB x Stimulus² interactions, $\beta = 1.0718$, 95% CI [1.0168 – 1.1298], $\chi^2(1, N=10300) = 6.6471$, $p = 0.010$, were significant. These interactions were visually and statistically close to identical to the previously documented Distress-PB interactions in the model without the other personality variables, and therefore is not interpreted further in this section. Additionally, the Orderliness-PB x Stimulus², $\beta = 0.9582$, 95% CI [0.9190 – 0.9991], $\chi^2(1, N=10300) = 4.0104$, $p = 0.045$, interaction was significant. As can be seen in Figure 24, those higher on Orderliness-PB showed a shallower decline in responding from CS+ to GS3. However, the CIs for the fitted plot substantially overlap at this point of the stimulus dimension, potentially indicating that this pattern seen in the actual values is not reflected in the fitted values. All other interactions from this model and all interaction terms from the model which only includes linear interactions were not significant.

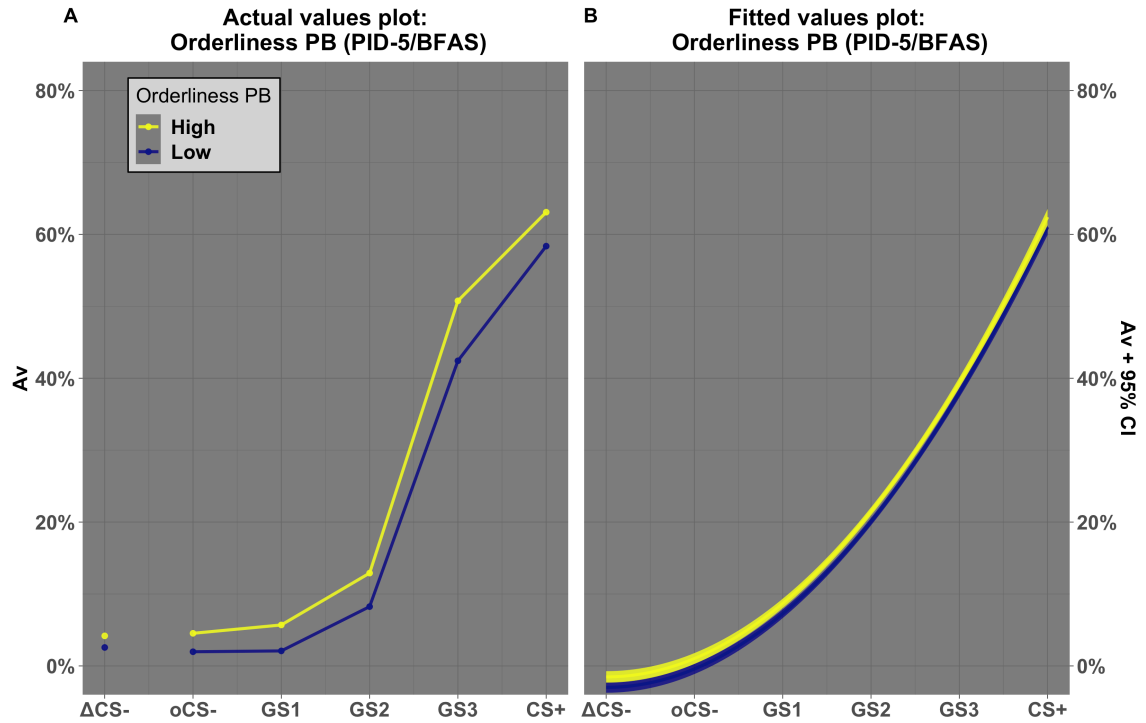


Figure 23. Visual representation of the Orderliness-PB x Stimulus² interaction while controlling for all other PB variables, with both actual avoidance % values (A) and fitted avoidance % (B) represented in separate plots. For graphing purposes, a median split was used to plot mean values for those with high or low Orderliness-PB scores. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine 95% confidence intervals (visible as semi-transparent borders around each slope). Av = avoidance; BFAS = Big Five Aspect Scale; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; PB = PID-5/BFAS composite; PID-5 = Personality Inventory for DSM-5.

Narrowband individual differences.

There was a significant main effect of TF-44, $\beta = 1.3352$, 95% CI [1.0805 – 1.6500], $\chi^2(1, N=10300) = 7.1641$, $p = 0.007$, and IUSF, $\beta = 1.2893$, 95% CI [1.0503 – 1.5827], $\chi^2(1, N=10300) = 5.9024$, $p = 0.015$, on avoidance; for both effects, increases in the narrowband trait corresponded to increased avoidance. The ASI and STAI-T main effects were not significant.

In the model with both linear and quadratic interaction terms, there were significant ASI x Stimulus $\beta = 0.7527$, 95% CI [0.5862 – 0.9666], $\chi^2(1, N=10300) =$

4.9573, $p = 0.026$, and ASI x Stimulus², $\beta = 1.0492$, 95% CI [1.0098 – 1.0901], $\chi^2(1, N=10300) = 6.0609$, $p = 0.014$, interactions. As seen in Figure 25, this appears to reflect a non-generalization effect, with those higher on ASI avoiding the CS+ and GS3 at a higher rate than those lower on ASI, but without a corresponding change in the grade of the decline from CS+ to GS3. This figure also shows a slightly shallower decline in responding from GS2 to GS1 in those with higher values of ASI. In a separate model that also included linear and quadratic interaction terms, there were significant STAI-T x Stimulus, $\beta = 0.7584$, 95% CI [0.5913 – 0.9728], $\chi^2(1, N=10300) = 4.7393$, $p = 0.029$, and STAI-T x Stimulus², $\beta = 1.0499$, 95% CI [1.0109 – 1.0903], $\chi^2(1, N=10300) = 6.3675$, $p = 0.012$, interactions. As with ASI and seen in Figure 25, this appears to reflect a non-generalization effect, with those higher on STAI-T avoiding the CS+ and GS3 at a higher rate than those lower on STAI-T, but without a corresponding change in the grade of the decline from CS+ to GS3. All other narrowband interactions were not significant.

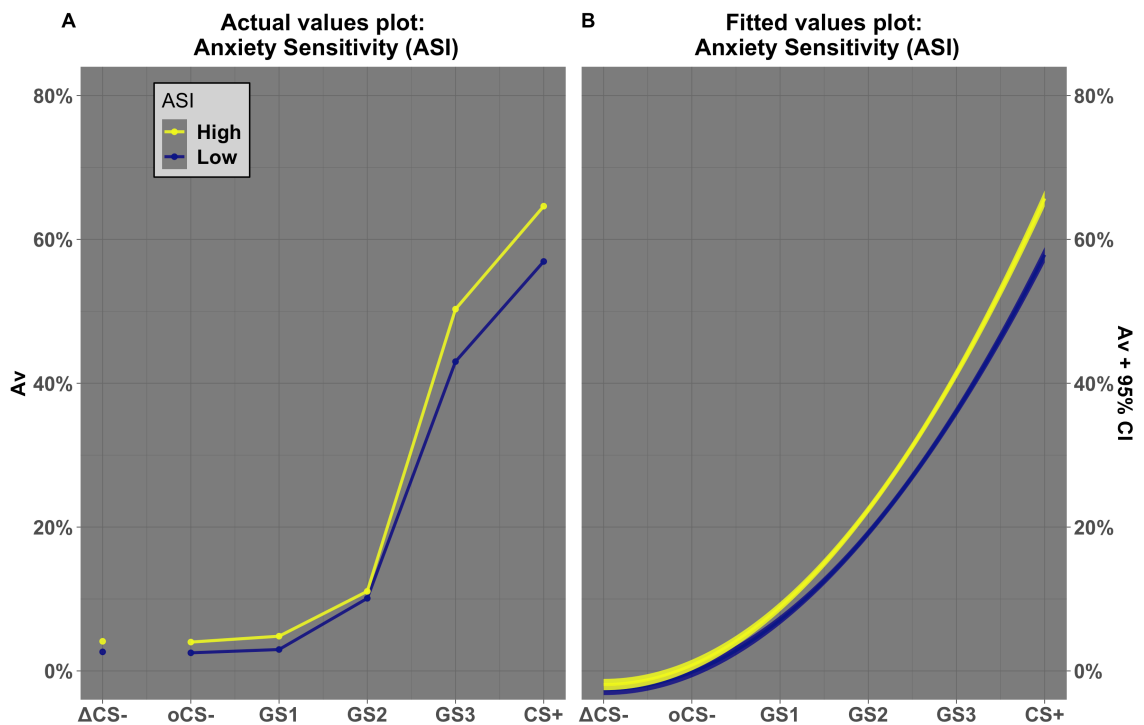


Figure 24. Visual representation of the ASI x Stimulus² interaction, with both actual avoidance % values (A) and fitted avoidance % (B) represented in separate plots. For graphing purposes, a median split was used to plot mean values for those with high or low ASI scores. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine 95% confidence intervals (visible as semi-transparent borders around each slope). ASI = Anxiety Sensitivity Inventory; Av = avoidance; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus.

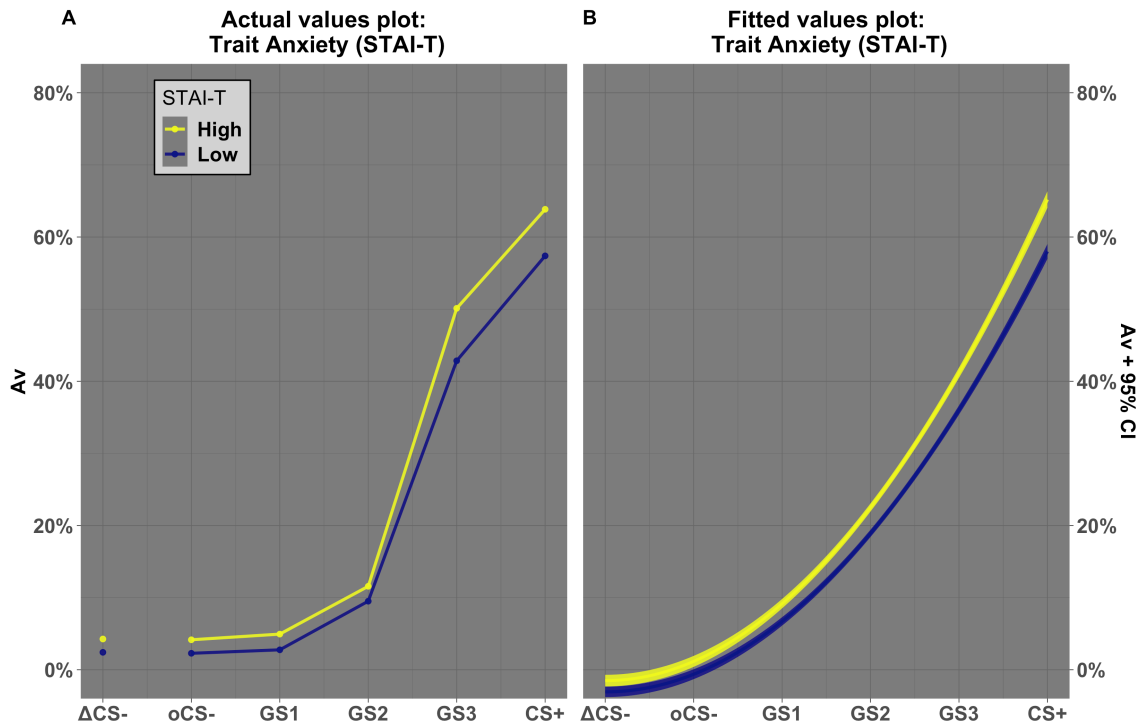


Figure 25. Visual representation of the STAI-T x Stimulus² interaction while controlling for all other PB variables, with both actual avoidance % values (A) and fitted avoidance % (B) represented in separate plots. For graphing purposes, a median split was used to plot mean values for those with high or low STAI-T scores. For the fitted values plot, a curvilinear smoothing function ($y \sim x + x^2$) was used to fit the best fitting curvilinear slope and to determine 95% confidence intervals (visible as semi-transparent borders around each slope). Av = avoidance; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; STAI-T = State Trait Anxiety Inventory – Trait.

Interim summary for Aim 1.

See Table 6 for an overview of significant main effects and interactions for the analyses conducted to address Aim 1. In terms of main effects, results were moderately consistent across risk rating and avoidance analyses: Negative Affectivity, TF-44, and

IUSF all predicted greater responding on each outcome measure. Distress-PB was also a consistent positive predictor, however, this is likely due to the influence of the Negative Affectivity variance that is part of the variable. Of the other personality variables primarily reflecting negative affect, Neuroticism was only found to be a predictor of overall risk ratings, and ASI and STAI-T did not predict overall levels of risk ratings or avoidance. Taken together, it appears that the more pathological extremes of negative affect are most consistently linked with increased threat appraisal and avoidance. The only personality variables with a negative association with an outcome variable were Assertiveness and Disinhibition, as both predicted lower levels of avoidance. Given the similar conceptual domains of these variables (i.e., tendency towards action and approach) these results might be indicators of a single determinant that contributes to decreased threat appraisal and avoidance. Notably, no main effects were found when predicting startle.

For interactions, especially as they pertain to potential generalization effects and associated personality traits, interpretation is more complicated. No personality variable significantly interacted with the Stimulus dimension (linear or quadratic) consistently across all three outcome measures, and thus we summarize interaction results sequentially by outcome variable. In terms of interactions predicting startle, both Extraversion and Assertiveness-PB (which is partially comprised of Extraversion variance) strengthened the effect of the Stimulus² predictor on startle. This can be interpreted as higher levels of Extraversion and Assertiveness-PB were associated with a stronger quadratic component, which corresponds to greater generalization for those lower on these dispositions. For both variables, this interaction appeared to be driven by a

shallower decline in startle from GS2 to GS1 for those with lower levels of the disposition; additionally, for Assertiveness-PB, lower levels were associated with a shallower decline from CS+ to GS3. The opposite relationship was observed for TF-44 and IUSF: both variables weakened the effect of Stimulus² predictor on startle, and therefore were associated with a weaker (more linear) quadratic relationship that corresponds to greater generalization. For both variables, this appears to be driven by shallower declines in startle potentiation from the CS+ to the GS1.

For interactions predicting risk ratings, Neuroticism, Negative Affectivity, Distress-PB (the composite comprised of items from both of the first two scales), and IUSF all strengthened the effect of the Stimulus (linear) predictor on risk ratings. This can be interpreted as higher levels of these variables were associated with a more linear decrease across the stimulus continuum (i.e., shallower decline), and therefore with greater generalization. For all of these variables, this interaction appeared to primarily be driven by a shallower decline in response from the CS+ to the GS3. Another personality predictor, Detachment, also strengthened the effect of the Stimulus predictor and was associated with greater generalization, but this particular interaction appeared to be driven by a shallower decline from GS2 to GS1, and not CS+ to GS3. Only one predictor was associated with less generalization of risk ratings: those higher on Enthusiasm-PB produced slightly more precipitous gradients than those lower on the variable.

There were some similarities in interaction results between risk rating and avoidance models. For interactions predicting avoidance, Negative Affectivity and Distress-PB also all predicted increased generalization, but in this case via an interaction with the Stimulus² predictor, not the Stimulus predictor. These interaction resulted in a

weakening of the quadratic effect in avoidance responding and appear to be driven by slightly shallower declines in avoidance from CS+ to GS3, as well as from GS2 to oCS- for Negative Affectivity only. Also similar to risk rating interaction results: Detachment predicted increased generalization, again via an interaction with the Stimulus² predictor that resulted in a weakening of the quadratic effect in avoidance responding. Dissimilarly to the Negative Affectivity/Distress-PB interactions, those higher on Detachment showed a shallower decline in avoidance from GS2 to oCS-; there was no observed difference in the CS+ to GS3 response slope. Additionally, and unique to prediction of avoid avoidance, Volatility, Conscientiousness, and Orderliness (a Conscientiousness aspect) weakened the quadratic effect in responding and therefore were associated with increased generalization. The same interaction was found for the Orderliness-PB variable. In all four interactions, those with higher levels of the personality variable demonstrated shallower declines in avoidance from the CS+ to GS3. Finally, significant Neuroticism x Stimulus², ASI x Stimulus² and STAI-T x Stimulus² interactions were found but based on the graphed gradients did not appear to reflect a generalization process.

Table 6. *Summary table for Aim 1: Significant Main Effects and Interactions by Personality Variable*

Personality Variable	Main Effects			Interactions		
	Startle	Risk Ratings	Avoidance	Startle	Risk Ratings	Avoidance
N		↑		↑ (L) ^b		↔ (Q) ^a
W						
V			↑ ^a			↑ (Q)
C						↑ (Q) ^b
In						
Or						↑ (Q)
E				↓ (Q) ^a		
As			↓ ^a			

En				
NA	↑ ^a	↑ ^a	↑ (L) ^a	↑ (Q) ^b
DE			↑ (L) ^a	↑ (Q) ^a
DI		↓ ^a		
DPB	↑ ^a	↑ ^a	↑ (L) ^a	↑ (Q) ^b
IPB				
OPB				↑ (Q) ^a
APB			↓ (Q) ^a	↑ (Q) ^a
EPB		↑ ^a	↓ (L) ^a	
TF-44	↑	↑	↑ (Q)	
IUSF	↑	↑	↑ (Q)	↑ (L)
ASI				↔ (Q)
STAI-T				↔ (Q)

Note: Only significant effects/interactions are explicitly documented in this table; a blank cell indicates a significant effect/interaction was not found. An upwards arrow (↑) indicates that the corresponding personality variable was positively associated with the dependent variable, or in the case of interactions, associated with increased generalization; the downwards arrow (↓) indicated the opposite (negative association/less generalization). A significant interaction marked with a left-right arrow (↔) indicates that although the interaction is significant, it does not appear to be driven by a generalization effect and the directionality is unclear. Interactions are marked as linear (L) or quadratic (Q) interactions, based on which model contained the significant interaction. All significant effects or interactions were $p < .05$. Composite traits refer to those derived from a combination of PID-5 and BFAS scales and are indicated with a suffix of "PB". As = Assertiveness; ASI = Anxiety Sensitivity Index; APB = Assertiveness PID-BFAS; C = Conscientiousness; DE = Detachment; DI = Disinhibition; DPB = Distress PID-BFAS; En = Enthusiasm; EPB = Enthusiasm PID-BFAS; E = Extraversion; In = Industriousness; IPB = Industriousness PID-BFAS; IUSF: Intolerance of Uncertainty – Short Form; OPB = Orderliness PID-BFAS; Or = Orderliness; N = Neuroticism; STAI-T: State Trait Anxiety Inventory – Trait; TF-44: Trait Fear 44-item; W = Withdrawal; V = Volatility.

^a significant in a model containing the other personality variables of the same type (i.e., BFAS, PID-5, or PB).

^b significant both when modeled separately and with other personality variables of the same type.

Aim 2 (A2): Testing improved prediction of avoidance in APIC models.

A2: Startle predicting avoidance.

Testing of the “base” APIC model that included startle as a fixed effect determined that a multilevel model continued to be appropriate for the data ($ICC = .32$). Initial model fitting resulted in the base model, Model 0 (random-intercept only, Stimulus modeled as a linear continuum only), as the only acceptable model for the following analyses based on our established criteria for APIC models; all models with a more complex random-effects structure (i.e., with at least one random slope) resulted in singular fits and were not appropriate for APIC analyses. The base model explained a significant proportion of variance in avoidance, $R_M^2 = .663$.

Normative broadband individual differences.

Neuroticism significantly improved model fit, $\chi^2(1) = 6.9754, p = 0.008, R_M^2 = .669$, as did the Volatility aspect, $\chi^2(1) = 8.9619, p = 0.001, R_M^2 = .671$, which was also a significant improvement on the Neuroticism model, $\chi^2(1) = 1.9865, p < 0.001$. Withdrawal did not significantly improve model fit. Conscientiousness did not significantly improve model fit, nor did the Industriousness and Orderliness aspects improve model fit when modeled separately. When modeled together, Industriousness and Orderliness significantly improved model fit, $\chi^2(1) = 7.4907, p = 0.026, R_M^2 = .670$. Extraversion did not significantly improve model fit, nor did the Assertiveness and Enthusiasm aspects improve model fit when modeled separately or together. The multi-trait model, which included Neuroticism, Conscientiousness, and Extraversion together, significantly improved model fit, $\chi^2(3) = 10.543, p = 0.014, R_M^2 = .672$. This model did not significantly improve on the Neuroticism-only model, which was the only single-trait model that represented an improvement over the base model, and therefore indicates that

the improved prediction of the multi-trait model is largely driven by the influence of Neuroticism. The multi-aspect model, which included all six aspects (Withdrawal, Volatility, Industriousness, Orderliness, Assertiveness, and Enthusiasm), significantly improved model fit, $\chi^2(6) = 19.064$, $p = 0.004$, $R_M^2 = .680$. This model significantly improved on all previous models ($ps \leq 0.036$), with the exception of the Volatility-only model, which the multi-aspect model did not significantly improve on.

Pathological broadband individual differences.

None of the models that included the pathological traits, both modeled individually and together, yielded significant improvements in model fit.

Combined PID-5/BFAS “full-spectrum” broadband individual differences (PB).

Distress-PB significantly improved model fit, $\chi^2(1) = 6.9754$, $p = 0.008$, $R_M^2 = .669$. None of the other composite variables (Industriousness-PB, Orderliness-PB, Assertiveness-PB, Enthusiasm-PB) significantly improved model fit when modeled separately. The multi-trait model for composite variables also did not significantly improve model fit.

Narrowband individual differences.

Three of the four tested narrowband individual differences significantly improved model fit when modeled separately: TF-44, $\chi^2(1) = 12.572$, $p < 0.001$, $R_M^2 = .671$, IUSF, $\chi^2(1) = 7.6594$, $p = 0.005$, $R_M^2 = .667$, and STAI-T, $\chi^2(1) = 7.7381$, $p = 0.005$, $R_M^2 = .669$, all significantly improved model fit; ASI did not significantly improve model fit. TF-44 represented a significant improvement in model fit over the models containing IU and

STAI-T separately; additionally, the STAI-T model significantly improved fit over the model containing the IU (all $ps < .001$).

A2: Risk ratings predicting avoidance.

Testing of the “base” APIC model that included risk ratings as a fixed effect determined that a multilevel model continued to be appropriate for the data ($ICC = .31$). As with the startle APIC model fitting, the initial model fitting resulted in the base model, Model 0 (random-intercept only, Stimulus modeled as a linear continuum only), as the only acceptable model for the following analyses based on our established criteria for APIC models; all models with a more complex random-effects structure (i.e., with at least one random slope) resulted in singular fits and were not appropriate for APIC analyses. The base model explained a significant proportion of variance in avoidance, $R_M^2 = .659$.

Normative broadband individual differences.

Neuroticism significantly improved model fit, $\chi^2(1) = 4.3462, p = 0.037, R_M^2 = .662$, as did the Volatility aspect, $\chi^2(1) = 6.243, p = 0.012, R_M^2 = .664$, which was also a significant improvement on the Neuroticism model, $\chi^2(1) = 1.9865, p < 0.001$. Withdrawal did not significantly improve model fit. Conscientiousness did not significantly improve model fit, nor did the Industriousness and Orderliness aspects improve model fit when modeled separately or together. Extraversion did not significantly improve model fit, nor did the Assertiveness and Enthusiasm aspects improve model fit when modeled separately or together. The multi-trait model, which included Neuroticism, Conscientiousness, and Extraversion together, did not significantly improve model fit. The multi-aspect model, which included all six aspects (Withdrawal, Volatility, Industriousness, Orderliness, Assertiveness, and Enthusiasm), significantly

improved model fit, $\chi^2(6) = 16.48, p = 0.011, R_M^2 = .672$. This model significantly improved on the Neuroticism-only model, $\chi^2(5) = 12.133, p = 0.033, R_M^2 = .672$, but did not significantly improve on the Volatility-only model.

Pathological broadband individual differences.

None of the models that included the pathological traits, both modeled individually and together, yielded significant improvements in model fit.

Combined PID-5/BFAS “full-spectrum” broadband individual differences (PB).

Assertiveness-PB significantly improved model fit, $\chi^2(1) = 4.0381, p = 0.044, R_M^2 = .663$. None of the other composite variables (Distress-PB, Industriousness-PB, Orderliness-PB, Enthusiasm-PB) significantly improved model fit when modeled separately. The multi-trait model for composite variables also did not significantly improve model fit.

Narrowband individual differences.

Three of the four tested narrowband individual differences significantly improved model fit when modeled separately: TF-44, $\chi^2(1) = 10.748, p = 0.001, R_M^2 = .666$, IUSF, $\chi^2(1) = 6.1403, p = 0.013, R_M^2 = .662$, and STAI-T, $\chi^2(1) = 5.398, p = 0.022, R_M^2 = .664$, all significantly improved model fit; ASI did not significantly improve model fit. TF-44 represented a significant improvement in model fit over the models containing IU and STAI-T separately; additionally, the IU model significantly improved fit over the model containing the IU (all $ps < .001$).

Interim summary for Aim 2.

See Table 7 for an overview of significant main effects and interactions for the analyses conducted to address Aim 2. Overall, results indicated that normative, but not pathological, broadband negative affect variables appeared to significantly improve prediction when included in APIC models. Unexpectedly, the Volatility aspect of Neuroticism appeared to drive Neuroticism-related model improvements. Further, there was limited evidence for Conscientiousness or Extraversion-related improvements, with the only significant improvements related to Industriousness and Orderliness (for APIC models including startle) or Assertiveness-PB (for APIC models including risk ratings). The strongest improvements for broadband normative traits for both startle and risk ratings models were associated with the multi-aspect models. Narrowband results were largely in line with predictions, with a primary finding that TF-44 is associated with the largest improvements in model fit, and STAI-T, but not ASI, unexpectedly improving model fit. Consistency between significant results for startle and risk rating APIC models was moderately strong, with the majority of the observed inconsistencies related to models involving aspect-level traits and composite variables.

Table 7. *Summary table for Aim 2: Significant Improvements in APIC Model Fit and Corresponding Effect Size (Marginal R^2).*

	Startle APIC (base $R_M^2 = .663$)	Risk Ratings APIC (base $R_M^2 = .659$)
Normative broadband (BFAS)		
N	0.669	0.662
W		
V	0.671	0.664
W, V		
C		
In		
Or		
In, Or	0.670	

E		
As		
En		
As, En		
Multi-trait	0.672	
Multi-aspect	0.680	0.672
Composite broadband (PB)		
DPB	0.669	
IPB		
OPB		
APB		0.663
EPB		
Multi-trait		
Narrowband		
TF-44	0.671	0.666
IUSF	0.667	0.662
ASI		
STAI-T	0.669	0.664

Note: Marginal R^2 is only displayed for models that significantly improved on the base model (i.e., the APIC model without any personality variables included); a blank cell indicates the corresponding model was not a significant improvement. Significant model improvements were $p < .05$. Bolded numbers indicate this was the highest effect size (i.e., best fit) among the type of variable (normative or composite) for either startle or risk rating models. Due to a lack of significant model improvement involving pathological (i.e., PID-5 variables) variables, they are not included in this table. Composite traits refer to those derived from a combination of PID-5 and BFAS scales and are indicated with a suffix of "PB". Multi-trait indicates a model with all of the trait level variables included (model with all normative traits or a model with all compound traits); multi-aspect is unique to normative models and indicates a model that includes all of the aspects in a single model. APIC = Aversive Pavlovian-Instrumental Covariation; As = Assertiveness; ASI = Anxiety Sensitivity Index; APB = Assertiveness PID-BFAS; BFAS = Big Five Aspect Scale; C = Conscientiousness; DE = Detachment; DI = Disinhibition; DPB = Distress PID-BFAS; En = Enthusiasm; EPB = Enthusiasm PID-BFAS; E = Extraversion; In = Industriousness; IPB = Industriousness PID-BFAS; IUSF: Intolerance of Uncertainty – Short Form; OPB = Orderliness PID-BFAS; Or = Orderliness; PID-5 = Personality Inventory for DSM-5; N = Neuroticism; R_M^2 = marginal R squared; STAI-T: State Trait Anxiety Inventory – Trait; TF-44: Trait Fear 44-item; W = Withdrawal; V = Volatility.

Aim 3 (A3): Moderation of APIC by dispositional traits.

Analyses for this aim were done using the same base APIC models used in Aim 2, which were random-intercept only models that only included the linear Stimulus term. This applied to both startle and risk ratings analyses. Due to a large number of null results, we streamline organization in the following sections to highlight significant moderators. Additionally, interpretation of significant interactions is limited to a broad overview of the interaction; more specific interpretations related to APIC-G vs APIC-CS+ and comparison between Pavlovian response variables are in the following sections as appropriate.

A3: Startle predicting avoidance.

We found a significant Orderliness x Startle x Stimulus effect, $\beta = 1.2431$, 95% CI [1.0214 – 1.5130], $\chi^2(1, N=1723) = 4.7145$, $p = 0.030$, which is evidence of moderation of APIC by Orderliness at specific levels of the Stimulus dimension. As can be seen in Figure 26, higher Orderliness generally facilitated the positive relationship between startle and avoidance for the CS+ to the GS1, and to the same degree across these stimuli (seen in the relatively consistent simple slope steepness). Most importantly, the difference in simple slopes for lower Orderliness highlights the significant moderation effect identified in the model: for the CS+, lower Orderliness weakened the association between startle and avoidance, such that higher startle was associated with *decreased* avoidance, whereas for the GS2 and GS1, lower Orderliness facilitated the association. For GS3, simple slopes appear to indicate a modest moderation effect, with a flatter lower Orderliness slope at this level resulting in a relatively stronger association between startle and avoidance for those higher on Orderliness. Taken together, it appears

that Orderliness differentially affects APIC across the stimulus continuum. All other tested moderation effects involving other personality variables were not significant.

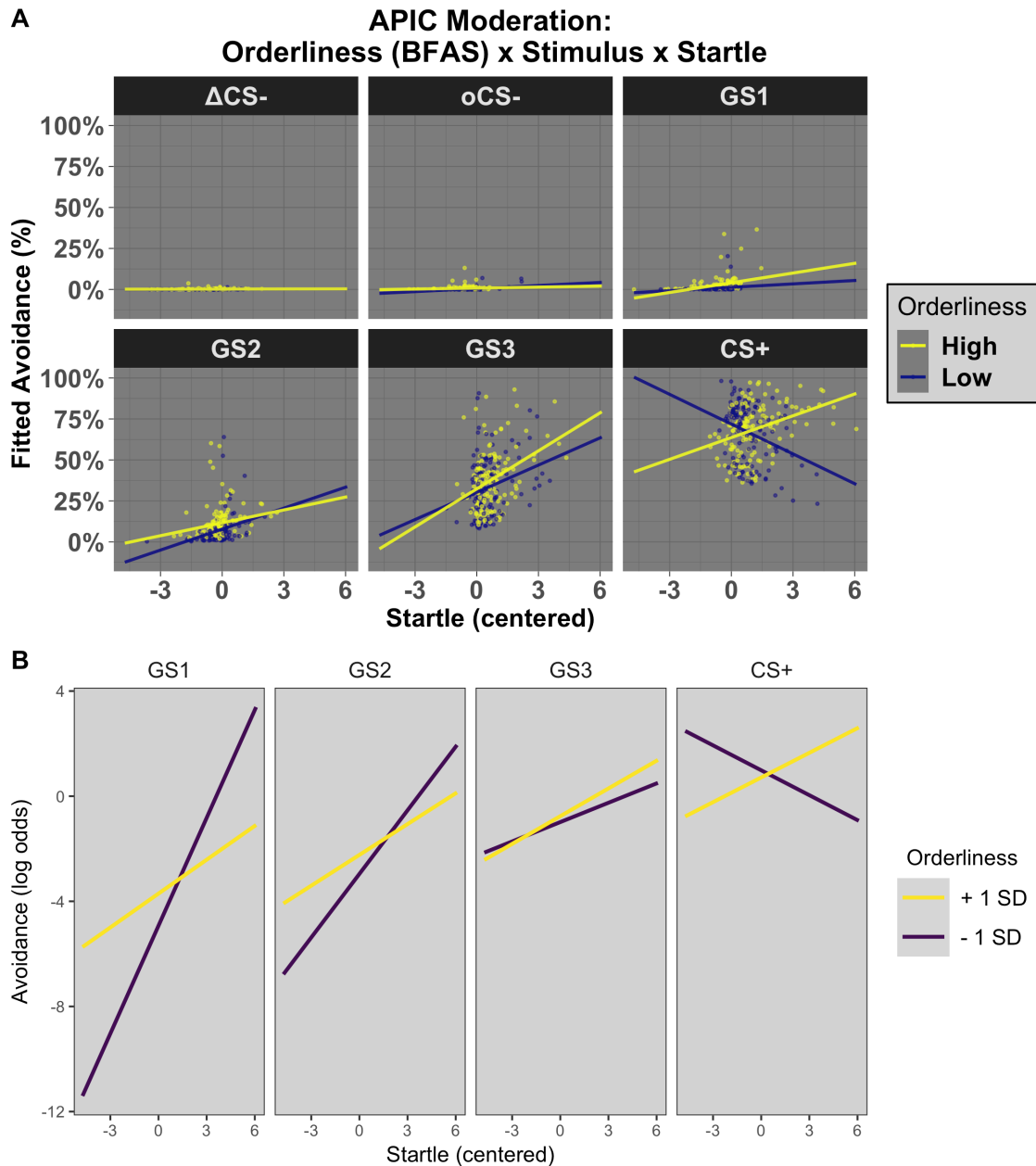


Figure 26. A: visual representation of the Orderliness x Stimulus x Startle interaction that represents a form of APIC moderation. Each subplot displays the association between fitted avoidance values and centered within cluster (CWC) startle at each Stimulus level. For graphing purposes, a median split was used to establish high and low Orderliness groups, and a linear smoothing function ($y \sim x$) was then used to approximate a best fitting summary line for each group, which are then transposed on the scatterplot showing individual data points. B: simple regression slopes depicting the moderating effect of

Orderliness at the level of the GSs and the CS+ at low (-1 SD) and high (+1 SD) levels of Orderliness. Avoidance is represented in log odds, such that the higher values represent an increased probability for avoidance of that associated stimulus. Simple slopes for each stimulus class are calculated separately, and therefore all avoidance values represent relative change (i.e., log odds values correspond to different absolute predicted avoidance rates for each stimulus, see A for absolute avoidance values). Slopes are extended to graph borders to help highlight the relative strength of the moderator effect (i.e., the interaction) for high/low levels of the moderator. APIC = Aversive Pavlovian Instrumental Covariation; BFAS = Big Five Aspect Inventory; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus.

A3: Risk ratings predicting avoidance.

We found four significant moderation effects for risk rating APIC analyses; all significant results were three-way interactions that are evidence of moderation of APIC by the tested personality variable at specific levels of the Stimulus dimension. There was a significant Conscientiousness x Stimulus x Risk Rating interaction, $\beta = 1.2431$, 95% CI [1.0214 – 1.5130], $\chi^2(1, N=1723) = 4.7145$, $p = 0.030$, as well as a significant Orderliness x Stimulus x Risk Rating interaction, $\beta = 1.3486$, 95% CI [1.0536 – 1.7264], $\chi^2(1, N=1728) = 5.6368$, $p = 0.018$. As can be seen in Figure 28, higher Conscientiousness facilitated the positive relationship between risk ratings and avoidance at the level of the CS+, whereas lower Conscientiousness facilitated this relationship at the level of the GS2 and GS1; this same pattern was also seen with Orderliness (see Figure 29), likely indicating that the moderating effect of Conscientiousness is driven by the Orderliness aspect. Also of note for these interactions is the discrepancy between the fitted values and simple slopes plots for the GS1; this appears to be driven by a relative outlier which, although not unduly influential on the model (i.e., does not affect statistical results when removed), biases the visualized summary fit line, and therefore simple slopes are primarily used for interpretation.

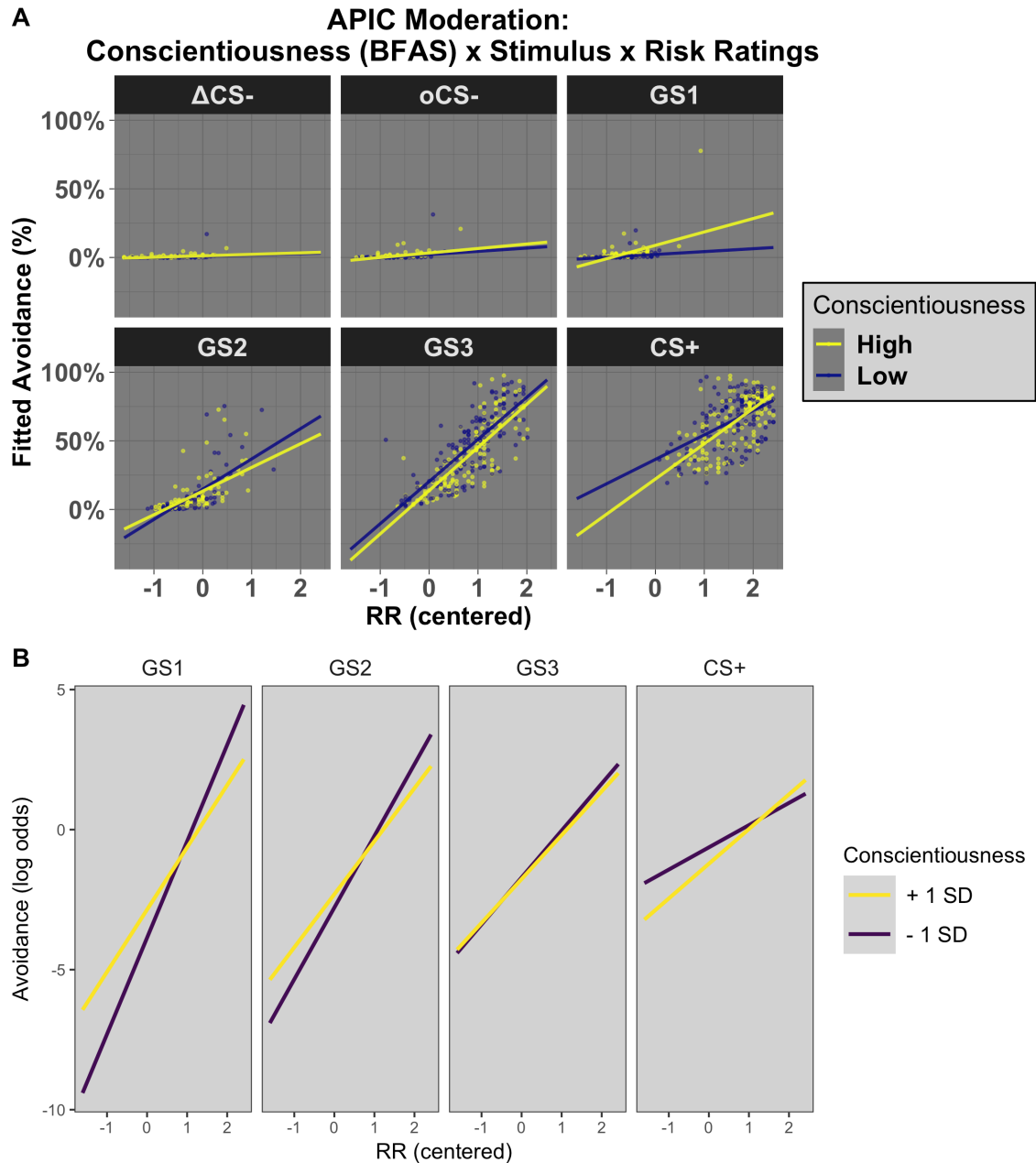


Figure 27. A: visual representation of the Conscientiousness x Stimulus x Risk Rating interaction that represents a form of APIC moderation. Each subplot displays the association between fitted avoidance values and centered within cluster (CWC) risk ratings at each Stimulus level. For graphing purposes, a median split was used to establish high and low Conscientiousness groups, and a linear smoothing function ($y \sim x$) was then used to approximate a best fitting summary line for each group, which are then transposed on the scatterplot showing individual data points. B: simple regression slopes depicting the moderating effect of Conscientiousness on risk ratings at the level of the GSs and the CS+ at low (-1 SD) and high (+1 SD) levels of Conscientiousness. Avoidance is represented in log odds, such that the higher values represent an increased probability for avoidance of that associated stimulus. Simple slopes for each stimulus

class are calculated separately, and therefore all avoidance values represent relative change (i.e., log odds values correspond to different absolute predicted avoidance rates for each stimulus, see A for absolute avoidance values). Slopes are extended to graph borders to help highlight the relative strength of the moderator effect (i.e., the interaction) for high/low levels of the moderator. APIC = Aversive Pavlovian Instrumental Covariation; BFAS = Big Five Aspect Inventory; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; RR = risk rating.

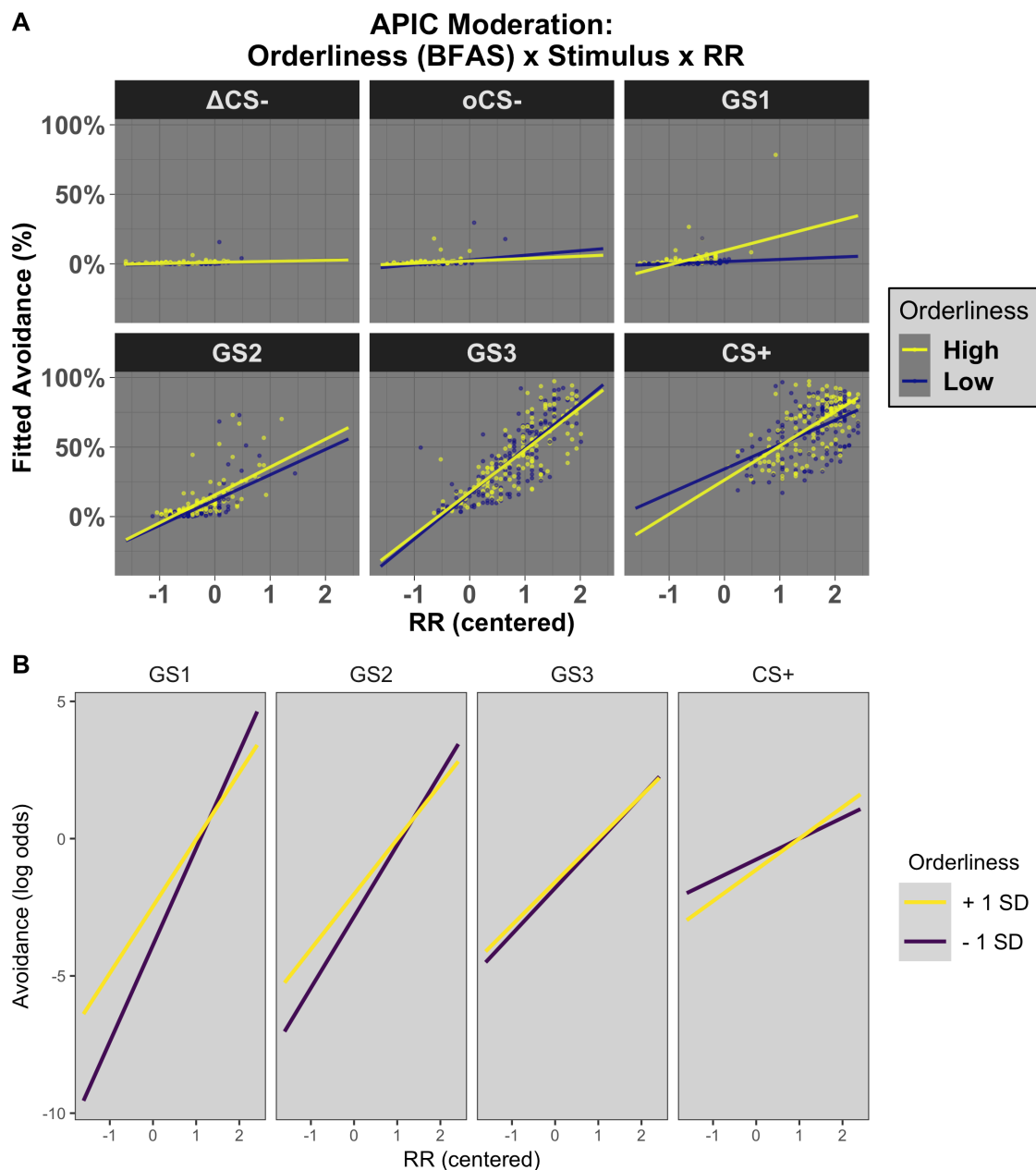


Figure 28. A: visual representation of the Orderliness x Stimulus x Risk Rating interaction that represents a form of APIC moderation. Each subplot displays the

association between fitted avoidance values and centered within cluster (CWC) risk ratings at each Stimulus level. For graphing purposes, a median split was used to establish high and low Orderliness groups, and a linear smoothing function ($y \sim x$) was then used to approximate a best fitting summary line for each group, which are then transposed on the scatterplot showing individual data points. B: simple regression slopes depicting the moderating effect of Orderliness on risk ratings at the level of the GSs and the CS+ at low (-1 SD) and high (+1 SD) levels of Orderliness. Avoidance is represented in log odds, such that the higher values represent an increased probability for avoidance of that associated stimulus. Simple slopes for each stimulus class are calculated separately, and therefore all avoidance values represent relative change (i.e., log odds values correspond to different absolute predicted avoidance rates for each stimulus, see A for absolute avoidance values). Slopes are extended to graph borders to help highlight the relative strength of the moderator effect (i.e., the interaction) for high/low levels of the moderator. APIC = Aversive Pavlovian Instrumental Covariation; BFAS = Big Five Aspect Inventory; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; RR = risk rating.

There was also a significant Extraversion x Risk Rating x Stimulus interaction, $\beta = 1.3479$, 95% CI [1.0490 – 1.7321], $\chi^2(1, N=1728) = 5.4456$, $p = 0.020$, as well as a significant Assertiveness x Risk Rating x Stimulus interaction, $\beta = 1.3138$, 95% CI [1.0166 – 1.6979], $\chi^2(1, N=1728) = 4.3513$, $p = 0.037$. As can be seen in Figure 30, lower Extraversion facilitated the positive relationship between risk ratings and avoidance for all GSs; this facilitation effect appeared to strengthen as the GS became more dissimilar from the CS+ (i.e., strongest for the GS1, weakest for the GS3); this same pattern was also seen with Assertiveness (see Figure 31), likely indicating that the moderating effect of Extraversion is driven by the Assertiveness aspect. Also of note for these interactions is the discrepancy between the fitted values and simple slopes plots for the GS1; this appears to be driven by the same relative outlier from the Conscientiousness/Orderliness plots.

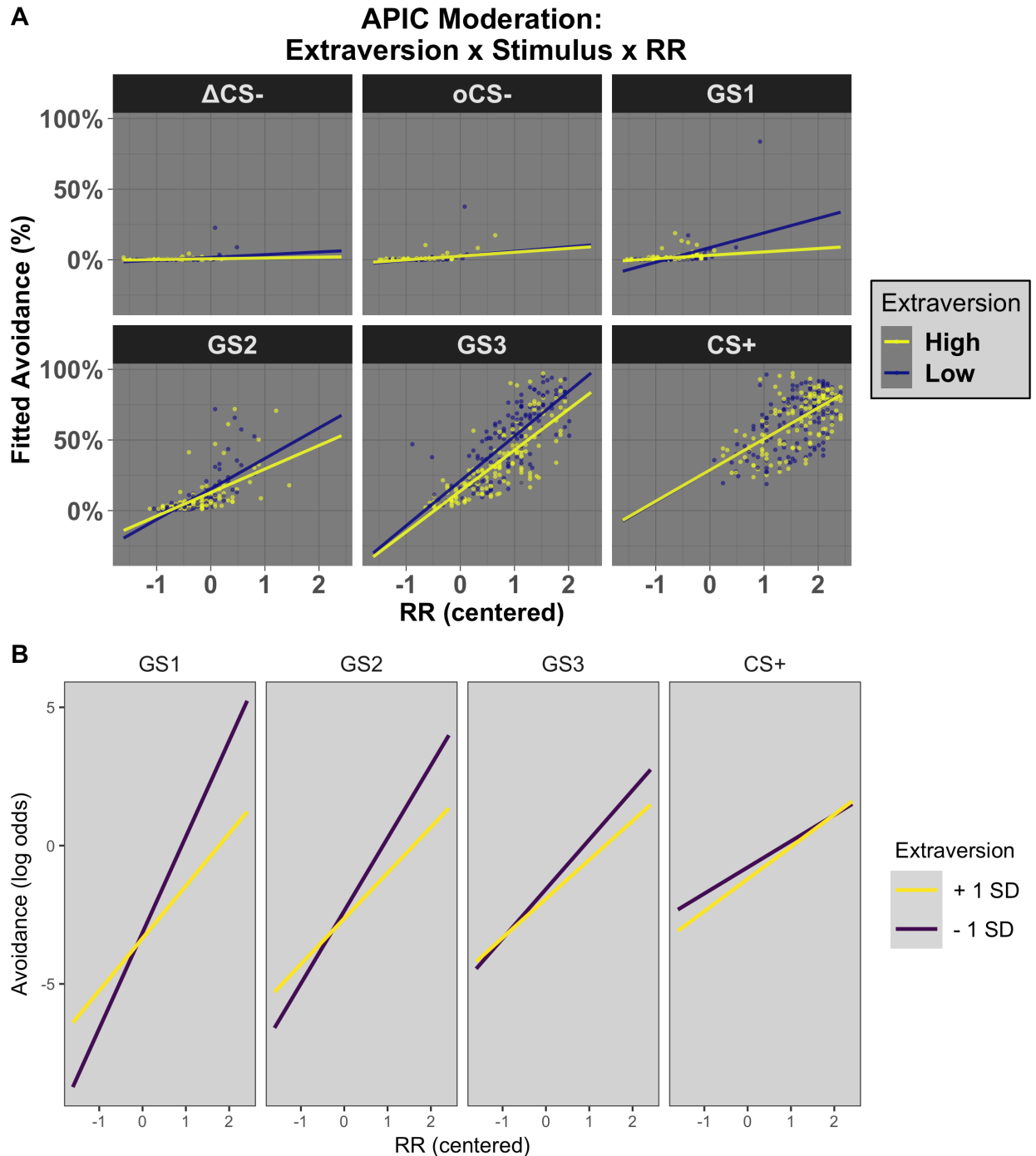


Figure 29. A: visual representation of the Extraversion x Stimulus x Risk Rating interaction that represents a form of APIC moderation. Each subplot displays the association between fitted avoidance values and centered within cluster (CWC) risk ratings at each Stimulus level. For graphing purposes, a median split was used to establish high and low Extraversion groups, and a linear smoothing function ($y \sim x$) was then used to approximate a best fitting summary line for each group, which are then transposed on the scatterplot showing individual data points. B: simple regression slopes depicting the moderating effect of Extraversion on risk ratings at the level of the GSs and

the CS+ at low (-1 SD) and high (+1 SD) levels of Extraversion. Avoidance is represented in log odds, such that the higher values represent an increased probability for avoidance of that associated stimulus. Simple slopes for each stimulus class are calculated separately, and therefore all avoidance values represent relative change (i.e., log odds values correspond to different absolute predicted avoidance rates for each stimulus, see A for absolute avoidance values). Slopes are extended to graph borders to help highlight the relative strength of the moderator effect (i.e., the interaction) for high/low levels of the moderator. APIC = Aversive Pavlovian Instrumental Covariation; BFAS = Big Five Aspect Inventory; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; RR = risk rating.

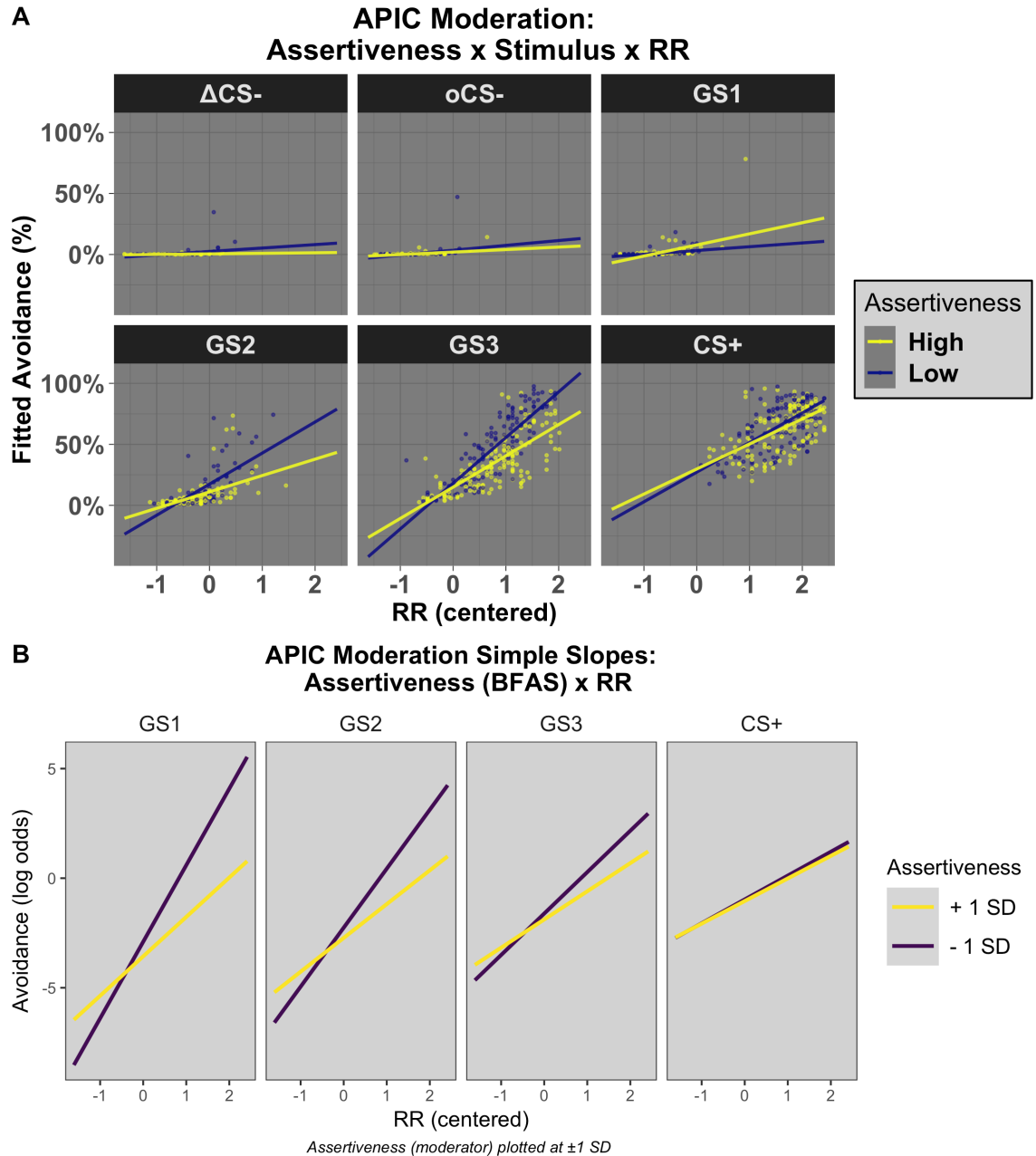


Figure 30. A: visual representation of the Assertiveness x Stimulus x Risk Rating interaction that represents a form of APIC moderation. Each subplot displays the association between fitted avoidance values and centered within cluster (CWC) risk ratings at each Stimulus level. For graphing purposes, a median split was used to establish high and low Assertiveness groups, and a linear smoothing function ($y \sim x$) was then used to approximate a best fitting summary line for each group, which are then transposed on the scatterplot showing individual data points. B: simple regression slopes depicting the moderating effect of Assertiveness on risk ratings at the level of the GSs and the CS+ at low (-1 SD) and high ($+1$ SD) levels of Assertiveness. Avoidance is represented in log odds, such that the higher values represent an increased probability for

avoidance of that associated stimulus. Simple slopes for each stimulus class are calculated separately, and therefore all avoidance values represent relative change (i.e., log odds values correspond to different absolute predicted avoidance rates for each stimulus, see A for absolute avoidance values). Slopes are extended to graph borders to help highlight the relative strength of the moderator effect (i.e., the interaction) for high/low levels of the moderator. APIC = Aversive Pavlovian Instrumental Covariation; BFAS = Big Five Aspect Inventory; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; RR = risk rating.

Interim summary for Aim 3.

When analyzing moderation of APIC broadly (i.e., not reducing the models to only test APIC-G or APIC-CS+), only normative traits functioned as significant moderators (see Table 8 for a summary of results from this section). Orderliness was the only personality dimension to moderate the APIC relationship for both risk ratings and startle, and in both cases lower levels of the aspect facilitated the positive relationship between the Pavlovian variable and avoidance at the level of the GS2 and GS1. Additionally, for both APIC models, higher levels of Orderliness facilitated the relationship at the level of the CS+. There were two differences between the two models: 1) higher Orderliness appeared to moderate startle, but not risk ratings, at the level of the GS3 and 2) lower Orderliness facilitated a *negative* relationship between startle and avoidance at the level of the CS+, potentially indicating a non-linear moderation effect of Orderliness for the CS+.

The remaining significant moderators were all moderators of the relationship between risk ratings and avoidance. Conscientiousness had a near identical pattern of moderation as Orderliness, with higher Conscientiousness facilitating the positive relationship between risk ratings at the level of the CS+, and lower Conscientiousness facilitating the relationship at GS2 and GS1. Finally, higher Extraversion and its aspect Assertiveness also demonstrated a near identical pattern of moderation, as lower levels of

both personality variables facilitated the positive relationship between risk ratings and avoidance for all GSs, with this moderating effect strengthening as the GS became more dissimilar from the CS+. Also of note for these interactions is the discrepancy between the fitted values and simple slopes plots for the GS1, which appeared to be driven by a relative outlier that did not affect the statistical models but biased the summary fit line. This potentially indicates that the method of dichotomization and summarization of the moderating variable had an outsized impact on visualization of these interactions, and that the observed GS1 moderation effects should be interpreted with caution.

We did not find significant APIC moderation effects for any other normative variables, or pathological or composite personality variables. There were also no significant moderation effects for narrowband personality traits.

Table 8. *Summary table for Aim 3: Significant APIC Moderators*

Personality Variable	Startle				Risk Ratings			
	GS1	GS2	GS3	CS+	GS1	GS2	GS3	CS+
N								
W								
V								
C					↓	↓	-	↑
In								
Or	↓	↓	↑	↑↓	↓	↓	-	↑
E					↓	↓	↓	-
As					↓	↓	↓	-
En								

Note: Only significant moderating effects are documented in this table, blank cells indicate a significant moderation effect was not found for that particular variable and stimulus pair. As only variables from the BFAS were found to moderate APIC, only BFAS variables are included. An upwards arrow (↑) indicates that higher levels of the corresponding personality variable facilitated the APIC relationship (i.e., potentiated the positive relationship); the downwards arrow (↓) indicated that lower levels of the corresponding personality variable facilitated the APIC relationship. A dash (-) in a cell indicates that although the corresponding personality variable was found to overall be a

moderator, it did not appear to moderate APIC at this level of the stimulus dimension. Both arrows in the same cell ($\uparrow\downarrow$) indicated that both high and low levels of the corresponding personality trait moderated the relationship, but in different directions (i.e., one facilitated the positive relationship, one attenuated the relationship) and is indicative of a non-linear moderation effect. All significant three-way interactions that indicated a moderation effect were $p < .05$. APIC = Aversive Pavlovian-Instrumental Covariation; As = Assertiveness; BFAS = Big Five Aspect Scale; C = Conscientiousness; En = Enthusiasm; E = Extraversion; In = Industriousness; Or = Orderliness; N = Neuroticism; W = Withdrawal; V = Volatility.

Aim 4 (A4): Specific APIC-G and APIC-CS+ effects.

Analyses for this aim were based on the same APIC models used in Aims 2 and 3, with modifications made to specifically test for APIC-G and APIC-CS+ as detailed in the Method section. This applied to both startle and risk ratings analyses. As we continued to use the basic APIC model for this aim we started to observe singular fits in our MLMs, which is due to these modified models using a reduced Stimulus dimension to probe APIC-G and CS+ separately and therefore having less within-subjects variation contributing to the model. Comparisons of these models to standard GLM models (i.e., logistic regressions) with the same model specifications (other than the random-intercept) revealed that the multilevel models yielded uniformly better model fits (lower Deviance and AIC statistics for generalized linear mixed models vs GLMs). We therefore continued to use an MLM approach for these analyses to allow for more straightforward interpretation of these results with results from other sets of analyses that use MLM. The overall ramifications of this issue on interpretation and our decision to continue using non-optimal multilevel models is reviewed in the Discussion section.

As with Aim 3, we streamline organization in the following sections to highlight significant moderators. We also highlight if significant moderators reflect a previously

identified moderation in the overall APIC analyses, or if they are unique to APIC-G/APIC-CS+ analyses.

A4: APIC-G: Startle predicting avoidance.

Notably, the Orderliness x Startle x Stimulus interaction seen in the overall startle APIC analyses was not significant for equivalent APIC-G analyses. Overall, we did not find any significant startle APIC-G moderators.

A4: APIC-G: Risk ratings predicting avoidance.

Of the significant risk rating APIC moderators, only Extraversion continued to be a significant moderator for risk rating APIC-G, as evidenced by an Extraversion x Risk Rating x Stimulus interaction, $\beta = 2.5523$, 95% CI [1.2275 – 5.3071], $\chi^2 (1, N=1152) = 6.2935$, $p = 0.012$. As can be seen in Figure 32, moderation of APIC-G was evident for the GS1 and GS2, but not GS3, with lower Extraversion facilitating the positive relationship between risk ratings and avoidance for these stimuli. Unique to APIC-G compared with APIC, the Enthusiasm x Stimulus x Risk Rating interaction was also significant, $\beta = 2.7104$, 95% CI [1.3179 – 5.5746], $\chi^2 (1, N=1152) = 7.3441$, $p = 0.007$, with the same pattern of moderation for the GSs as seen for Extraversion (see Figure 33). Additionally, higher Enthusiasm facilitated the relationship between risk ratings and avoidance for the GS3. As with the previous APIC analyses, a relative outlier biased the summary fit line, resulting in simple slopes being used as the primary method of interpretation.

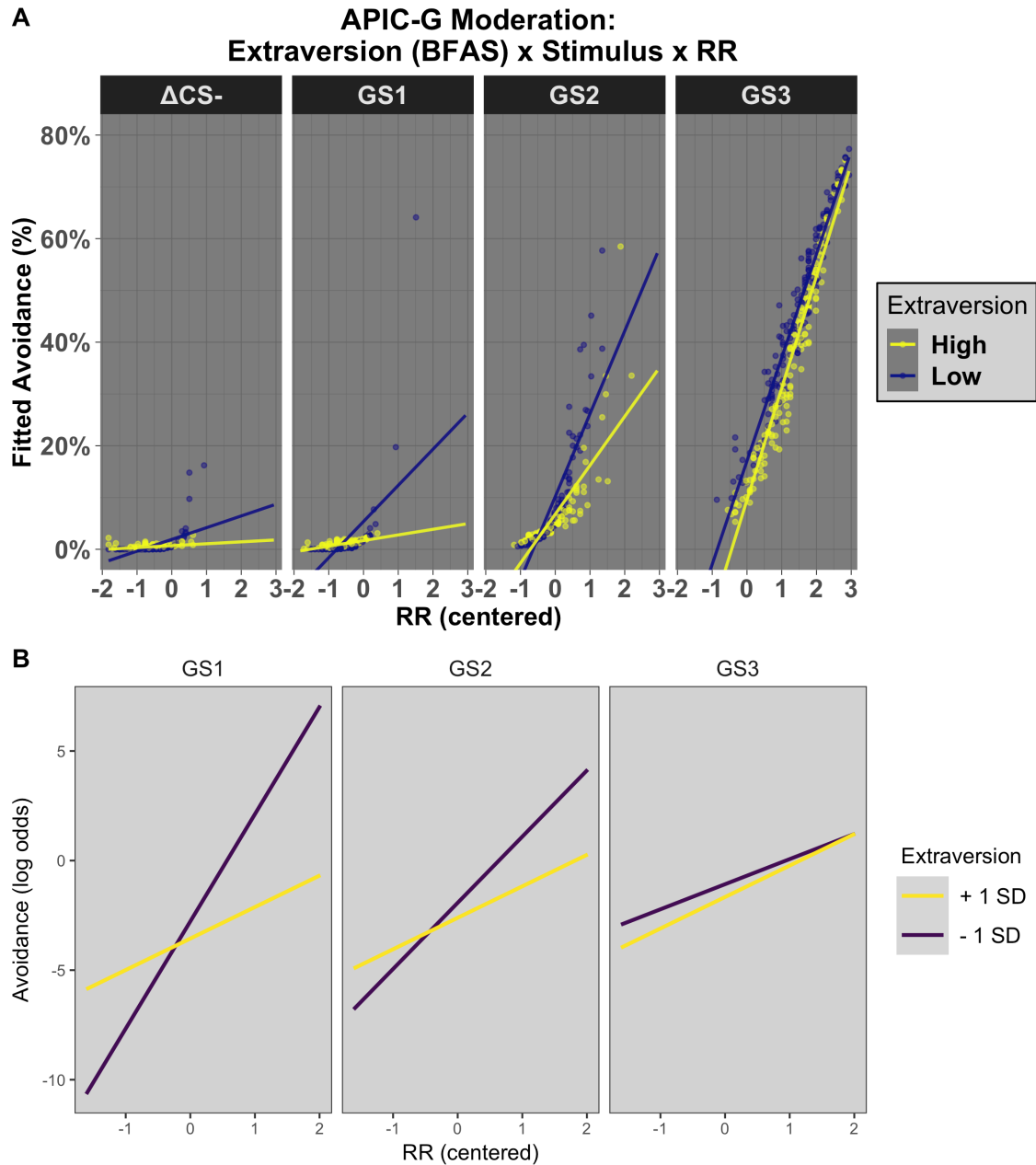


Figure 31. A: visual representation of the Extraversion x Stimulus x Risk Rating interaction that represents a form of APIC-G moderation. Each subplot displays the association between fitted avoidance values and centered within cluster (CWC) risk ratings at each Stimulus level. For graphing purposes, a median split was used to establish high and low Extraversion groups, and a linear smoothing function ($y \sim x$) was then used to approximate a best fitting summary line for each group, which are then transposed on the scatterplot showing individual data points. B: simple regression slopes depicting the moderating effect of Extraversion on risk ratings at the level of the GSs and the CS+ at low (-1 SD) and high (+1 SD) levels of Extraversion. Avoidance is represented in log odds, such that the higher values represent an increased probability for avoidance of that associated stimulus. Simple slopes for each stimulus class are

calculated separately, and therefore all avoidance values represent relative change (i.e., log odds values correspond to different absolute predicted avoidance rates for each stimulus, see A for absolute avoidance values). Slopes are extended to graph borders to help highlight the relative strength of the moderator effect (i.e., the interaction) for high/low levels of the moderator. APIC-G = Aversive Pavlovian Instrumental Covariation during Generalization; BFAS = Big Five Aspect Inventory; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; RR = risk rating.

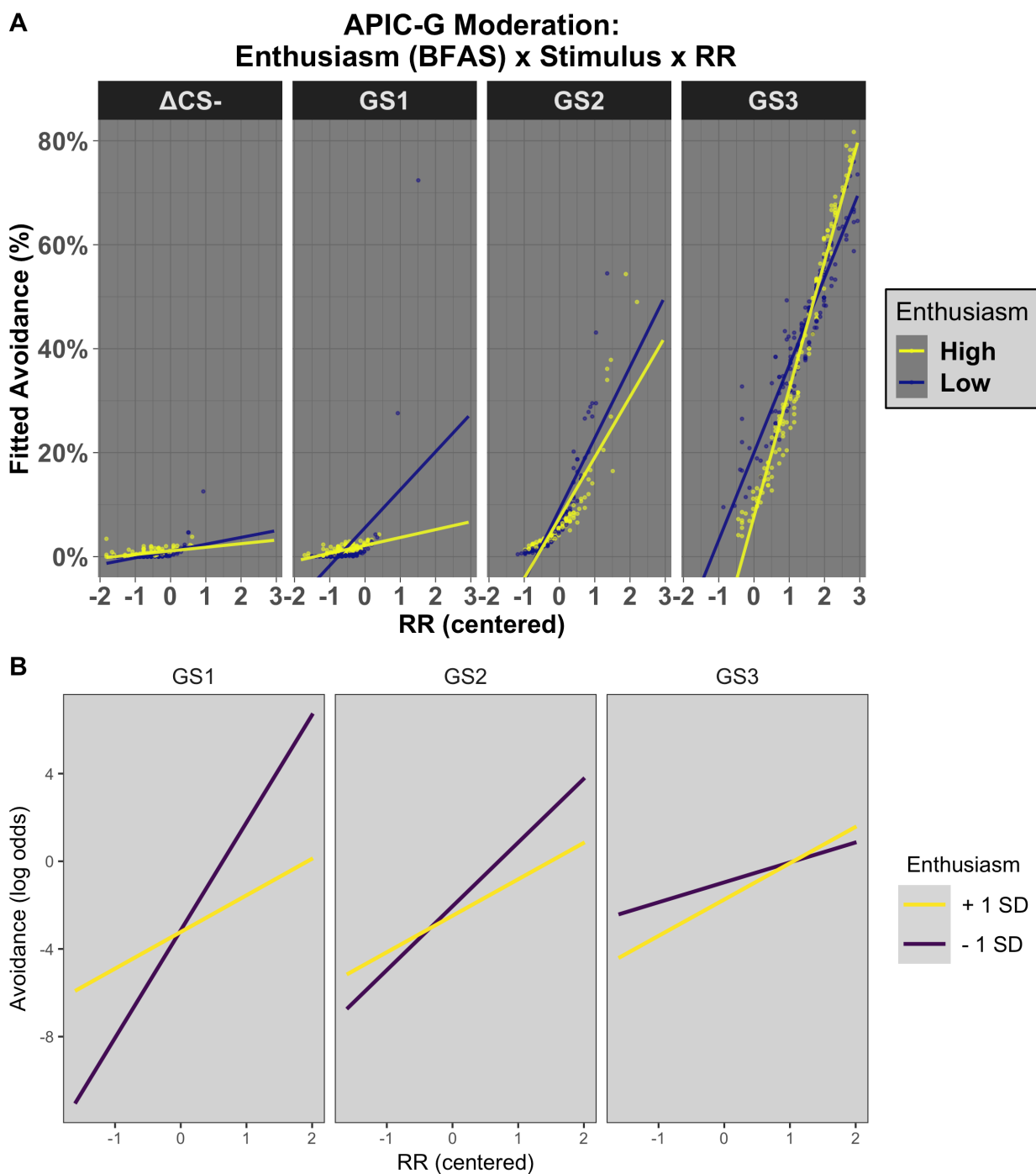


Figure 32. A: visual representation of the Enthusiasm x Stimulus x Risk Rating interaction that represents a form of APIC-G moderation. Each subplot displays the association between fitted avoidance values and centered within cluster (CWC) risk ratings at each Stimulus level. For graphing purposes, a median split was used to establish high and low Enthusiasm groups, and a linear smoothing function ($y \sim x$) was then used to approximate a best fitting summary line for each group, which are then transposed on the scatterplot showing individual data points. B: simple regression slopes depicting the moderating effect of Enthusiasm on risk ratings at the level of the GSs and the CS+ at low (-1 SD) and high (+1 SD) levels of Enthusiasm. Avoidance is represented in log odds, such that the higher values represent an increased probability for avoidance of that associated stimulus. Simple slopes for each stimulus class are calculated separately, and therefore all avoidance values represent relative change (i.e., log odds values correspond to different absolute predicted avoidance rates for each stimulus, see A for absolute avoidance values). Slopes are extended to graph borders to help highlight the relative strength of the moderator effect (i.e., the interaction) for high/low levels of the moderator. APIC-G = Aversive Pavlovian Instrumental Covariation during Generalization; BFAS = Big Five Aspect Inventory; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; RR = risk rating.

In terms of non-normative personality variables, Negative Affect was found to be a significant moderator, as evidenced by a Negative Affect x Stimulus x Risk Rating interaction, $\beta = 0.4436$, 95% CI [0.2229 – 0.8827], $\chi^2 (1, N=1152) = 5.3601$, $p = 0.021$. As can be seen in Figure 34, moderation of APIC-G was evident for the GS1 and GS2, with higher Negative Affectivity facilitating the positive relationship between risk ratings and avoidance for these stimuli. Additionally, in contrast to the other GSs, lower Negative Affectivity appears to facilitate the risk rating and avoidance relationship at the level of the GS3.

There was also evidence of Detachment functioning as an APIC moderator, with a significant Detachment x Stimulus x Risk Rating interaction, $\beta = 0.4749$, 95% CI [0.2317 – 0.9731], $\chi^2 (1, N=1152) = 4.1387$, $p = 0.042$, showing a near identical pattern of moderation as seen for Negative Affectivity (see Figure 35), with higher Detachment facilitating the positive relationship between risk ratings and avoidance for GS2 and GS1, but lower Detachment facilitating the relationship for the GS3.

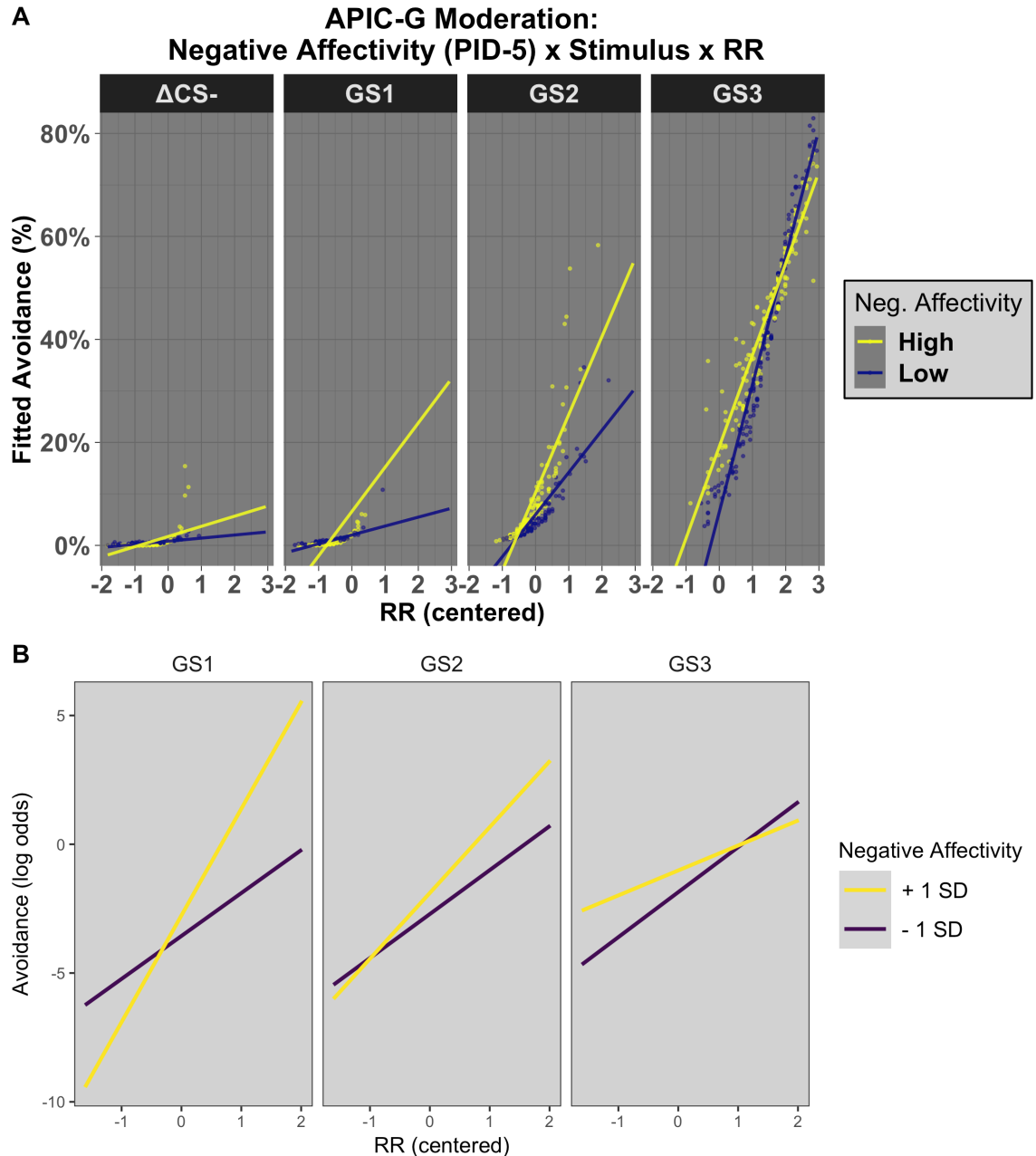


Figure 33. A: visual representation of the Negative Affectivity x Stimulus x Risk Rating interaction that represents a form of APIC-G moderation. Each subplot displays the association between fitted avoidance values and centered within cluster (CWC) risk ratings at each Stimulus level. For graphing purposes, a median split was used to establish high and low Negative Affectivity groups, and a linear smoothing function ($y \sim x$) was then used to approximate a best fitting summary line for each group, which are then transposed on the scatterplot showing individual data points. B: simple regression slopes depicting the moderating effect of Negative Affectivity on risk ratings at the level of the GSs and the CS+ at low (-1 SD) and high (+1 SD) levels of Negative Affectivity. Avoidance is represented in log odds, such that the higher values represent an increased

probability for avoidance of that associated stimulus. Simple slopes for each stimulus class are calculated separately, and therefore all avoidance values represent relative change (i.e., log odds values correspond to different absolute predicted avoidance rates for each stimulus, see A for absolute avoidance values). Slopes are extended to graph borders to help highlight the relative strength of the moderator effect (i.e., the interaction) for high/low levels of the moderator. APIC-G = Aversive Pavlovian Instrumental Covariation during Generalization; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; PID-5 = Personality Inventory for DSM-5; RR = risk rating.

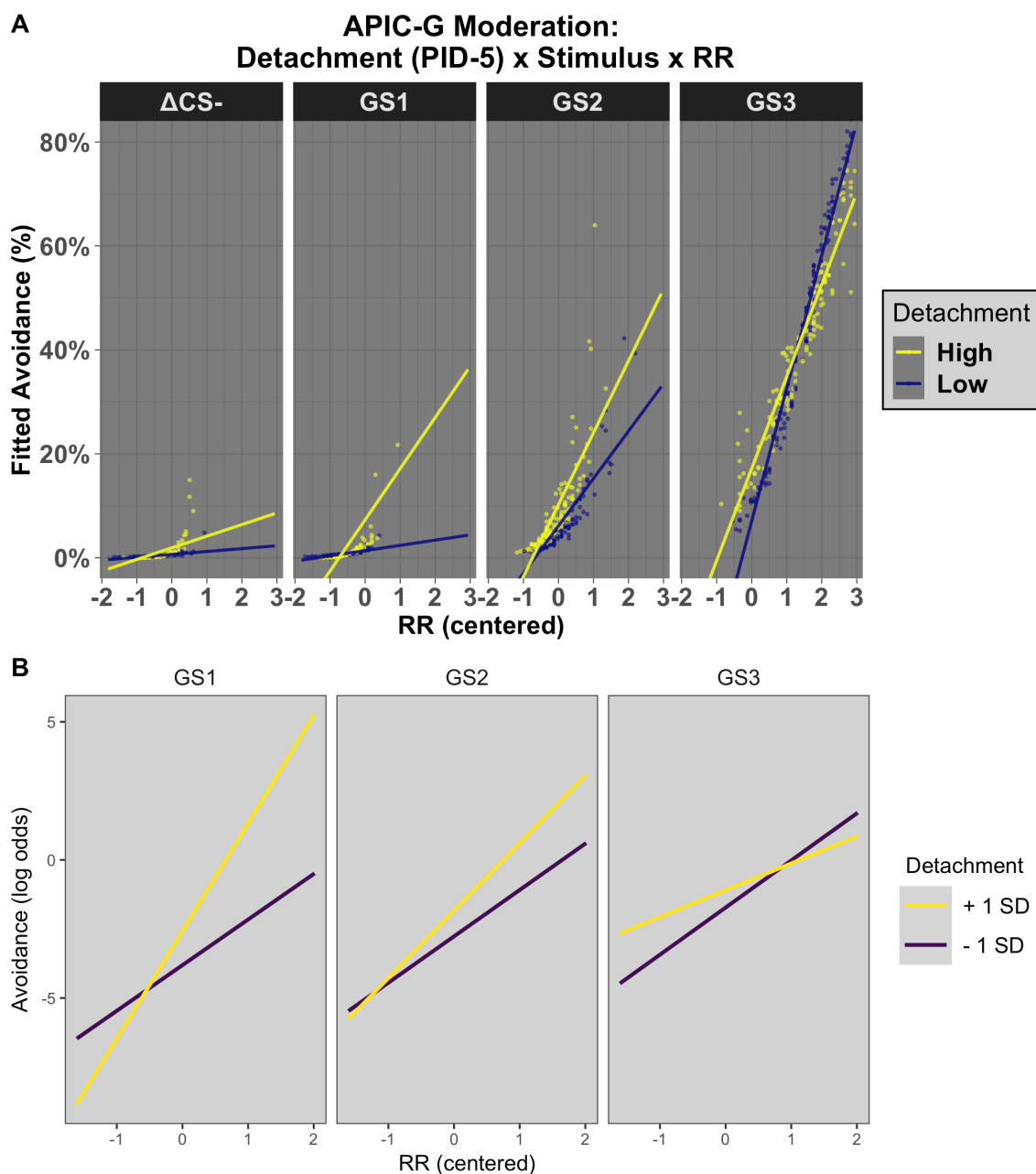


Figure 34. A: visual representation of the Detachment x Stimulus x Risk Rating interaction that represents a form of APIC-G moderation. Each subplot displays the association between fitted avoidance values and centered within cluster (CWC) risk ratings at each Stimulus level. For graphing purposes, a median split was used to establish high and low Detachment groups, and a linear smoothing function ($y \sim x$) was then used to approximate a best fitting summary line for each group, which are then transposed on the scatterplot showing individual data points. B: simple regression slopes depicting the moderating effect of Detachment on risk ratings at the level of the GSs and the CS+ at low (-1 SD) and high (+1 SD) levels of Detachment Avoidance is represented in log odds, such that the higher values represent an increased probability for avoidance of that associated stimulus. Simple slopes for each stimulus class are calculated separately, and therefore all avoidance values represent relative change (i.e., log odds values correspond to different absolute predicted avoidance rates for each stimulus, see A for absolute avoidance values). Slopes are extended to graph borders to help highlight the relative strength of the moderator effect (i.e., the interaction) for high/low levels of the moderator. APIC-G = Aversive Pavlovian Instrumental Covariation during Generalization; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; PID-5 = Personality Inventory for DSM-5; RR = risk rating.

Finally, two PID-5/BFAS composite variables were significant moderators. The Assertiveness-PB x Stimulus x Risk Ratings interaction was significant, $\beta = 2.4583$, 95% CI [1.2643 – 4.7800], $\chi^2 (1, N=1152) = 7.028$, $p = 0.008$. As can be seen in Figure 36, moderation of APIC-G was evident for the GS1 and GS2, with lower Assertiveness-PB facilitating the positive relationship between risk ratings and avoidance for these stimuli. The other PB variable related to Extraversion, Enthusiasm-PB, was also a significant moderator, as the Enthusiasm-PB x Stimulus x Risk Ratings interaction was significant, $\beta = 2.7104$, 95% CI [1.3179 – 5.5746], $\chi^2 (1, N=1152) = 7.3451$, $p = 0.007$. As can be seen in Figures 37, the moderating effect of Enthusiasm-PB resembles that of Assertiveness-PB at the level of the GS1, with lower Enthusiasm-PB facilitating the positive relationship between risk ratings and avoidance for these stimuli. However, in contrast to Assertiveness-PB, the facilitatory effect of Enthusiasm-PB at the level of the GS1 is less clear; the simple slope plot shows a slight moderation effect due to the lower-level slope, but this is not reflected in the summary fit lines shown in the equivalent fitted values plot,

potentially indicating the simple slopes plot is reflective of a moderating effect only at extremely low levels of Enthusiasm-PB. Also in contrast to Assertiveness-PB, Enthusiasm-PB appears to moderate the GS3 relationship, with higher levels of Enthusiasm-PB facilitating the risk rating and avoidance relationship for this stimulus.

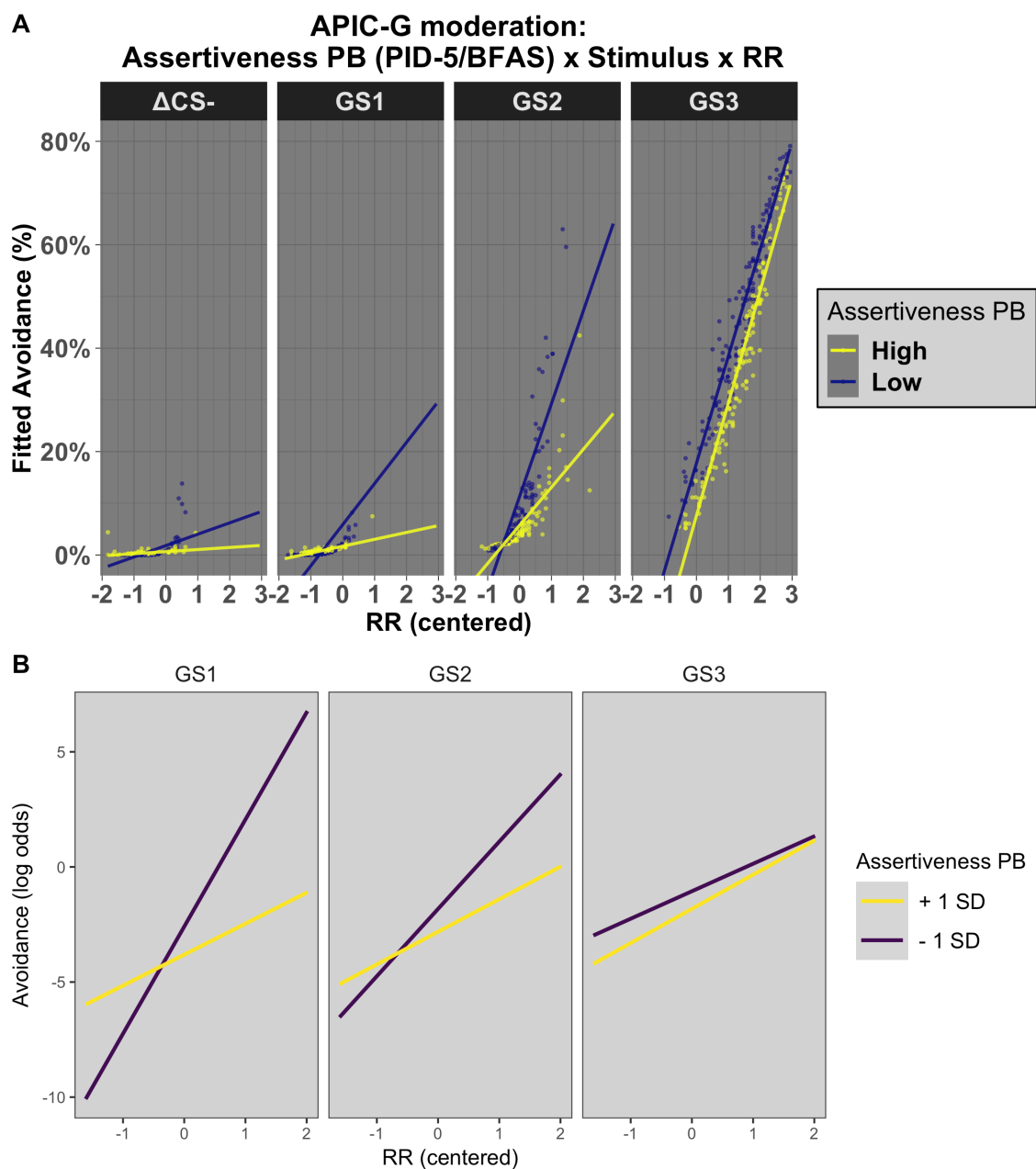


Figure 35. A: visual representation of the Assertiveness-PB x Stimulus x Risk Rating interaction that represents a form of APIC-G moderation. Each subplot displays the association between fitted avoidance values and centered within cluster (CWC) risk

ratings at each Stimulus level. For graphing purposes, a median split was used to establish high and low Assertiveness-PB groups, and a linear smoothing function ($y \sim x$) was then used to approximate a best fitting summary line for each group, which are then transposed on the scatterplot showing individual data points. B: simple regression slopes depicting the moderating effect of Assertiveness-PB on risk ratings at the level of the GSs and the CS+ at low (-1 SD) and high (+1 SD) levels of Assertiveness-PB. Avoidance is represented in log odds, such that the higher values represent an increased probability for avoidance of that associated stimulus. Simple slopes for each stimulus class are calculated separately, and therefore all avoidance values represent relative change (i.e., log odds values correspond to different absolute predicted avoidance rates for each stimulus, see A for absolute avoidance values). Slopes are extended to graph borders to help highlight the relative strength of the moderator effect (i.e., the interaction) for high/low levels of the moderator. APIC-G = Aversive Pavlovian Instrumental Covariation during Generalization; BFAS = Big Five Aspect Scale; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; PB = PID-5/BFAS composite; PID-5 = Personality Inventory for DSM-5; RR = risk rating.

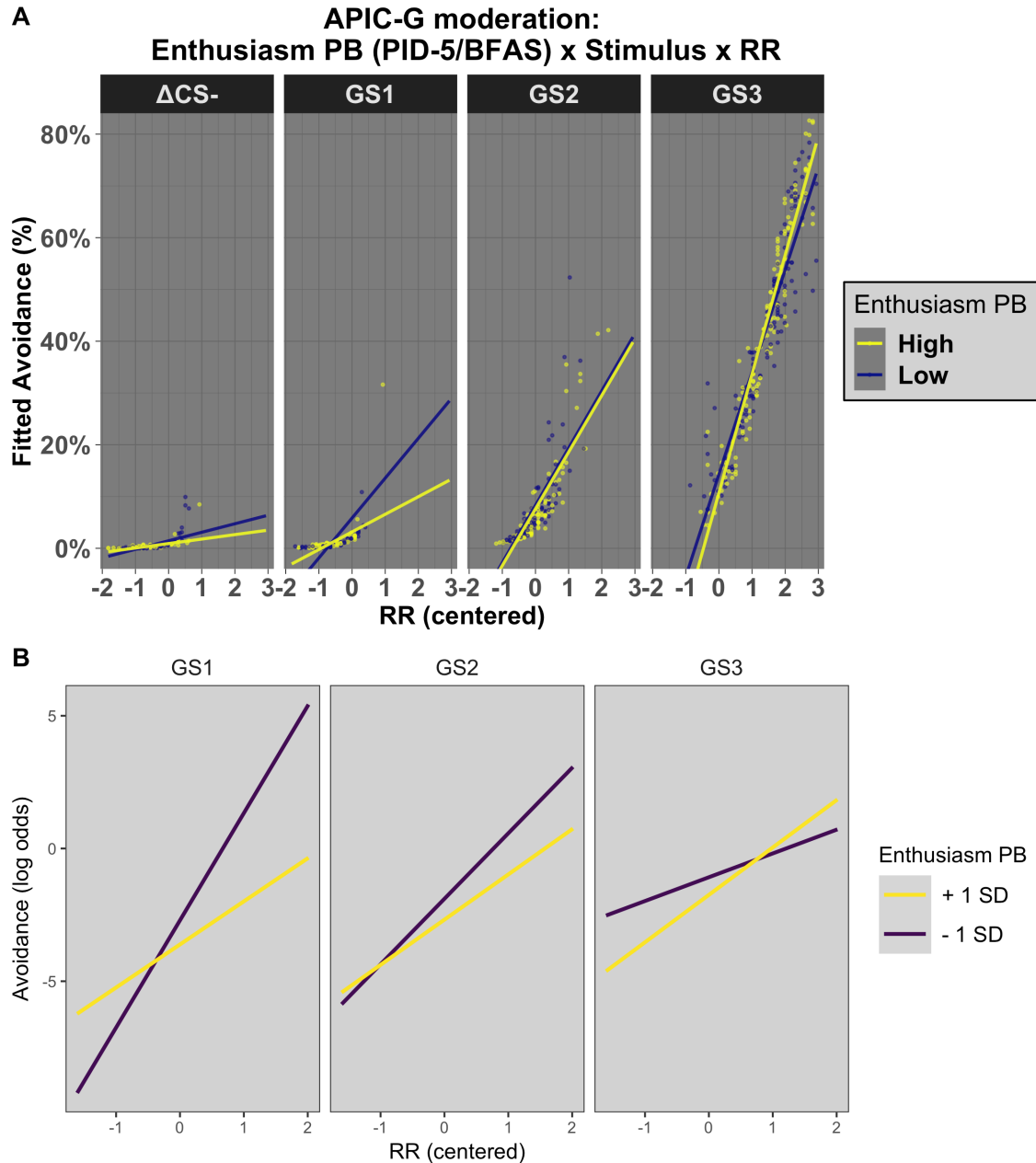


Figure 36. A: visual representation of the Enthusiasm-PB x Stimulus x Risk Rating interaction that represents a form of APIC-G moderation. Each subplot displays the association between fitted avoidance values and centered within cluster (CWC) risk ratings at each Stimulus level. For graphing purposes, a median split was used to establish high and low Enthusiasm-PB groups, and a linear smoothing function ($y \sim x$) was then used to approximate a best fitting summary line for each group, which are then transposed on the scatterplot showing individual data points. B: simple regression slopes depicting the moderating effect of Enthusiasm-PB on risk ratings at the level of the GSs and the CS+ at low (-1 SD) and high (+1 SD) levels of Enthusiasm-PB. Avoidance is represented in log odds, such that the higher values represent an increased probability for avoidance of that associated stimulus. Simple slopes for each stimulus class are

calculated separately, and therefore all avoidance values represent relative change (i.e., log odds values correspond to different absolute predicted avoidance rates for each stimulus, see A for absolute avoidance values). Slopes are extended to graph borders to help highlight the relative strength of the moderator effect (i.e., the interaction) for high/low levels of the moderator. APIC-G = Aversive Pavlovian Instrumental Covariation during Generalization; BFAS = Big Five Aspect Scale; CS- = conditioned safety cue; CS+ = conditioned danger cue; GS = generalization stimulus; PB = PID-5/BFAS composite; PID-5 = Personality Inventory for DSM-5; RR = risk rating.

A4: APIC-CS+: Startle predicting avoidance

Although there was a clear moderation effect for Orderliness on startle at the level of the CS+ for APIC analyses, we did not find a significant APIC-CS+ moderation effect for Orderliness. Overall, we did not find any significant risk rating APIC-CS+ moderators.

A4: APIC-CS+: Risk ratings predicting avoidance.

As with the startle models, we did not find any significant APIC-CS+ moderators for risk rating models. These null findings are despite clear moderation effects for Conscientiousness and Orderliness on risk ratings at the level of the CS+ for APIC analyses, as well as a less clear, but still notable, moderation effect for Assertiveness on risk ratings at the level of the CS+ for APIC analyses.

Interim summary for Aim 4.

In line with APIC analyses, we also found significant normative personality moderators of APIC-G; however, in contrast to APIC analyses we also found significant pathological and composite personality moderators (see Table 9 for summary of significant moderation effects for APIC-G). Additionally, all identified moderators were related to risk ratings. The only identified moderator of APIC that was also found for APIC-G was Extraversion. For APIC-G, lowered Extraversion facilitated the positive relationship between risk ratings and avoidance for the GS1 and GS2, but not the GS3.

This differed from the Extraversion moderation effect observed for APIC, which also included a facilitative effect of lower Extraversion at the level of the GS3. Additionally, the Assertiveness facet of Extraversion was found to be a significant APIC moderator, but was not a significant APIC-G moderator. Instead, Enthusiasm, the other Extraversion aspect, was a significant APIC-G moderator and demonstrated close to the same pattern of moderation as Extraversion. Additionally, and differing from Extraversion, higher Enthusiasm facilitated the relationship between risk ratings and avoidance for the GS3. This appears to be due to an overall steeper slope for those higher on Enthusiasm across the Stimulus continuum, which resulted in a more distinct moderation effect for the GS3 compared with the same effect for Extraversion.

Extraversion and Enthusiasm were the only normative personality variables to significantly moderate APIC-G. Negative Affectivity and Detachment were identified as significant pathological personality moderators of APIC-G, with both demonstrating a facilitating effect of the risk rating and avoidance relationship for the GS1 and GS2 at higher levels of the trait and for GS3 at lower levels of the trait. In terms of composite variables, both of the variables related to Extraversion, Assertiveness-PB and Enthusiasm-PB, were significant APIC-G moderators, with lower levels of both variables facilitating the risk rating and avoidance relationship at the level of the GS1, lower levels of Assertiveness-PB (and potentially Enthusiasm-PB, but there is some discrepancy in the summary fit lines versus the simple slopes) facilitating the relationship at the level of the GS2, and higher levels of Enthusiasm-PB facilitating the relationship at the level of the GS3.

Despite notable moderation effects observed at the level of the CS+ for APIC analyses, formal APIC-CS+ analyses (i.e., models only including the Δ CS- and CS+) did not identify any significant moderators. We also did not find significant moderation effects for any other broadband personality variables. There were also no significant moderation effects for narrowband personality traits.

Table 9. *Summary table for Aim 4: Significant APIC-G Moderators*

Personality Variable	Risk Ratings		
	GS1	GS2	GS3
N			
W			
V			
C			
In			
Or	-	↑	↓
E	↓	↓	-
As			
En	↓	↓	↑
NA	↑	↑	↓
DE	↑	↑	↓
DI			
DPB			
IPB			
OPB			
APB	↓	↓	-
EPB	↓	↓	↑

Note: Only significant moderating effects are documented in this table, blank cells indicate a significant moderation effect was not found for that particular variable and stimulus pair. As only broadband personality variables were found to be significant APIC-G moderators, and only for models involving risk ratings, narrowband variables and startle models are excluded from this table. An upwards arrow (↑) indicates that higher levels of the corresponding personality variable facilitated the APIC relationship (i.e., potentiated the positive relationship); the downwards arrow (↓) indicated that lower levels of the corresponding personality variable facilitated the APIC-G relationship. A

dash (-) in a cell indicates that although the corresponding personality variable was found to overall be a moderator, it did not appear to moderate APIC-G at this level of the stimulus dimension. All significant three-way interactions that indicated a moderation effect were $p < .05$. APIC-G = Aversive Pavlovian-Instrumental Covariation during Generalization; As = Assertiveness, APB = Assertiveness PID-BFAS, C = Conscientiousness, DE = Detachment, DI = Disinhibition, DPB = Distress PID-BFAS, En = Enthusiasm, EPB = Enthusiasm PID-BFAS, E = Extraversion, In = Industriousness, IPB = Industriousness PID-BFAS, OPB = Orderliness PID-BFAS, Or = Orderliness, N = Neuroticism, W = Withdrawal, V = Volatility.

Discussion

Overall, this study replicated the results of prior work from our group using the PIG paradigm (van Meurs et al., 2014) which found both Pavlovian and instrumental avoidance generalization effects. Crucially, we expanded on these findings by rigorously testing the relations between a broad range of personality individual differences and these generalization effects, and then extend this work by testing how these personality variables potentially improve prediction of and, in certain cases, moderate the covariation between Pavlovian fear and instrumental avoidance. Given the large number of analyses and the complexity of the results, we next present in-depth interpretations of results organized by the relevant aim and then the associated set of hypotheses, as well as an additional section discussing unaddressed results that were not related to specific hypotheses. These interpretations are then followed by a synthesis of all results.

Aim 1: Identifying personality predictors of generalization.

Overall, analyses for Aim 1 identified a number of predictors of generalization across different outcome measures. One notable pattern across the analyses for this aim was the lack of physiological generalization findings (i.e., when operationalized with fear-potentiated startle). Despite the relatively large sample size used for these analyses, it is not surprising that the fewest significant effects were found for startle, as the EMG

measurement of the eyeblink that is used to index startle is statistically noisier than behavioral measurements, and inconsistent results across personality traits were observed in the only comparable study that used startle eyeblink in relation to personality traits in a large sample (Gazendam et al., 2014). We address this issue in the overall synthesis of results that follows the more specific discussion for each aim, and instead focus on interpreting results for Aim 1 in psychological terms.

A1.H1.

Contrary to predictions, Withdrawal (a Neuroticism aspect) was not associated with greater generalization in any outcome measures. However, Neuroticism was associated with greater threat estimation generalization (as indexed by a shallower decline in risk ratings from CS+ to GS3), and Volatility (the other Neuroticism aspect) was associated with increased avoidance generalization (as indexed by a shallower decline in avoidance decisions from CS+ to GS3) when the other aspects of interest were held constant (i.e., included in the same model). These results provide tentative support for a role of higher, but normative, levels of negative affect in generalization. Neuroticism's relationship with increased threat estimation to the GS3 resembles that seen in studies of disorders associated with extreme levels of Neuroticism, such as PTSD (e.g., Kaczurkin et al., 2016) or GAD (e.g., Lissek et al., 2014), but with a notably weaker generalization effect (i.e., a steeper decline from CS+ to GS3 than that seen in the anxiety disorder groups). The Volatility generalization effect is also modest, but an appropriate comparison study is not available due to the lack of research related to avoidance generalization. Overall, it is not clear why we did not find Withdrawal effects, but instead found a Neuroticism and Volatility effect for different outcomes. One

possibility is that both aspects of Neuroticism are determinants of Pavlovian generalization, despite Withdrawal nominally being the “anxiety” aspect, and that when modeling the aspects separately their predictive power was weakened. Another possibility is that the item content in Withdrawal related to depression is as prominent as the anxiety-related content, and thus attenuates the relationship between the anxiety-related variance and Pavlovian generalization. There is some support for this in the current results, as Withdrawal is positively associated with a state measure of depression used in this study, $r = .55$. This does not help explain why Volatility predicted avoidance generalization, however. From one point of view, the finding appears quite contradictory to expectations given that Volatility is conceptualized as an aspect representing emotional lability and anger (DeYoung et al., 2007), especially considering the intuitive appeal of and empirical support for anger as the “approach” emotion (Carver & Harmon-Jones, 2009; Harmon-Jones, Peterson, Gable, & Harmon-Jones, 2008). One possibility is that the observed Volatility finding does not reflect a direct anger effect, and that PIG task is not activating an in-moment experience of anger in those higher on Volatility and therefore not explicitly activating approach motivation, but that the finding is instead capturing a tendency towards frustration in a subset of high-Volatility participants that find the reward to not be equivalent to the effort needed to obtain it (e.g., the act of winning is not worth the shock) and disengage to downregulate their activation. This tendency might then be most activated during the (likely) most challenging situation in the task: determining if the GS3 is dangerous. This is plausible given evidence that frustration is associated with avoidance as well as approach (e.g., Adelman & Maatsch, 1956; McNaughton, DeYoung, & Corr, 2016), but to avoid overinterpretation we

consider this to be an explanation that perhaps cannot be further clarified or supported using the data from this study.

A1.H2.

Our hypothesis that Negative Affectivity and Distress-PB (a Neuroticism + Negative Affectivity composite) would be associated with greater generalization was supported; both traits significantly predicted generalized threat evaluation and avoidance when accounting for the effects of the other variables of the same type (i.e., Detachment and Disinhibition for Negative Affectivity, the other PB variables for Distress-PB). The explanation that these variables are approximations of the disorders previously associated with generalization appears valid, especially given that the gradients resemble those findings from samples with pathology, but to a more modest degree. It is also important to note that these do not necessarily reflect separate findings, given that Distress-PB is partially comprised of Negative Affectivity items, and perhaps overall provides initial support for the utility of dimensional models in the study of generalization, as both pathological and “full-spectrum” variables were associated with generalization.

A1.H3.

Our hypothesis that Conscientious and its two aspects, Industriousness and Orderliness, would be differentially associated with avoidance generalization was partially supported: Orderliness was associated with increased avoidance generalization, whereas Industriousness did not have a significant effect. Further, Conscientiousness reflected the Orderliness effect and was also significantly associated with increased avoidance generalization. The association between Orderliness and generalization as indexed by shallower declines from the CS+ to the GS3 is understandable when viewing

the GS3 as an ambiguous approach-avoidance conflict, and the CS+ as a less ambiguous situation. Someone higher on Orderliness might not necessarily avoid the CS+ more than those lower on the aspect (which is reflected in our results) because in less-ambiguous approach-avoidance situations, the optimal outcome is clearer and behavior is less determined by individual differences (i.e., a strong situation; Lissek, Pine, & Grillon, 2006). However, during GS3 trials, Orderliness might promote increased avoidance because of a tendency towards closely following the rules, even when maladaptive (Hewitt & Flett, 2007), and implicitly constructing an avoidance “rule” for the GS3 based on its increased similarity to the CS+. Put another way, if the rule for the CS+ is to avoid, then that rule generalizes to the GS3 for those higher on Orderliness. It should also be noted that the pathological extension of Orderliness, the Orderliness-PB composite trait, significantly predicts generalization in a similar pattern to Orderliness, but to perhaps an even greater degree. However, we refrain from continued interpretation of this interaction due to overlapping CIs for the Orderliness-PB fitted lines.

Given that the Conscientiousness results resemble those for Orderliness, it is likely that the variance associated with Orderliness is the driving predictive element of the trait, and that Industriousness (i.e., the Conscientiousness variance that isn’t accounted for by Orderliness) has little effect on generalization. There is also the possibility that our manipulation is not well suited to elicit Industriousness effects, as the contingencies and decision-making aspects of the PIG are fairly simple, and Industriousness is perhaps more reflective of how people persevere in service of a distal goal or stay focused during a complex task (e.g., Hickman, Stromme, & Lippman, 1998).

A1.H4.

Our hypotheses for Extraversion and its aspect, Assertiveness, to be associated with decreased avoidance generalization were not supported: although lower Assertiveness was associated with overall increased levels of avoidance, there was no evidence of this being related to generalization. This is potentially related to the relatively low value of the reward in the PIG (a “win”), as it is possible that this reward is not sufficient for those higher on Extraversion/Assertiveness to increase their approach during an ambiguous situation (i.e., GS3) above and beyond their baseline level of approach. For example, someone higher on Assertiveness might generally respond to any approach-avoidance challenge with an increased rate of approach regardless of the associated reward (i.e., a tendency towards increased agency and desire for achievement across situations; e.g., Depue & Collins, 1999), but would not necessarily approach at a relatively higher rate during risk unless the reward value supersedes the potential cost. Further, those higher on Assertiveness are not necessarily fearless or at an extremely low level of fear and other relevant negative affect traits that we would expect Assertiveness to clearly capture differences in generalization related to a lack of fear and anxiety. That said, Assertiveness (and the superordinate Extraversion trait) are generally anticorrelated with Neuroticism and fear measures (r s for Assertiveness correlations in the current study range from $-.23$ [Neuroticism] to $-.55$ [TF-44], also see DeYoung, Carey, Krueger, & Ross, 2016; Quilty, DeYoung, Oakman, & Bagby, 2014), so it is possible that Assertiveness genuinely tracks lower fear as well as higher overall approach, but that this not result in a significant association with experimental outcomes for another reason. For example, as previously noted, Extraversion and subordinate factors are not perfect analogues for approach motivation, or even the broader positive affect construct (Depue

& Collins, 1999; Quilty et al., 2014), and that the socially-relevant content included in most measures, including the BFAS, is an imperfect fit for the PIG task, which has an extremely limited social component in the form of the video game context involving a farmer. We also did not explicitly measure social aspects of motivation during the PIG (e.g., “how much was your decision to avoid the short road related to your desire to not see the farmer embarrassed or hurt?”), making a more in-depth analysis of this aspect of the task difficult, if not impossible.

A1.H5.

Our hypotheses regarding narrowband variables were partially supported: trait fear and IU were significant predictors of Pavlovian generalization, although only IU significantly predicted both physiological and behavioral (i.e., risk ratings) generalization. The physiological generalization gradients for trait fear and IU are perhaps the most reminiscent findings of those seen in generalization studies with anxiety patient samples, with higher trait fear/IU associated with an overall shallower decline in responding and a notably shallower decline from the CS+ to the GS3. The IU, but not trait fear, association with behavioral generalization is unexpected. One possible explanation is related to the specific trait fear measure used, the TF-44. This measure was constructed and optimized using a biobehavioral approach that emphasizes a biological criterion, in this case the startle reflex, which operationalized in a similar fashion to the method used for the current study (Kramer, Patrick, Krueger, & Gasperi, 2012; Kramer et al., 2019). The risk ratings used in the current study are, at best, an indirect measure of fear, and are better conceptualized as a cognitive appraisal of threat. Therefore, the TF-44 might not be optimized to predict differences in generalization as measured by risk

ratings. This is also consistent with conceptualizations of the trait fear construct, which revolve around defensive responding to acute threat. Conversely, it could also be argued that IU is particularly well-suited to predicting behavioral generalization as operationalized by risk ratings, which are essentially a form of predicting uncertain outcomes. Thus, those who associate greater uncertainty with more risk and distress (i.e., those higher on IU) are likely also those who would demonstrate increased generalization.

The lack of association between IU or trait fear with avoidance generalization, but significant main effects of both on avoidance (consistent with prior work on IU and instrumental avoidance, Flores, López, Vervliet, & Cobos, 2018), is also unexpected. The main effect, in which higher levels of the trait are associated with overall increased avoidance, is a logical extension of both traits: IU and trait fear both capture content related to avoidance and reduced approach, and those higher on these traits are likely those who report higher daily avoidance in response to threat and uncertainty. The same logic, however, would also suggest that those higher on these traits are more likely to generalize their avoidance to a greater a degree, yet the current results do not support this. A possible explanation is related to one of the overarching tenets of this study: approach-avoidance conflict has both fear and non-fear determinants, and maladaptive resolution of these conflicts is not solely determined by fear-related factors. Thus, it is possible that the variance captured by the IU and trait fear measure is too specific to the fear and anxiety domain and not to the other important factors that predict avoidance generalization. This also might help explain contradictory fear conditioning results in the IU literature (e.g., J. T.-H Chen & Lovibond, 2016; Chin, Nelson, Jackson, & Hajcak, 2016; Nelson,

Weinberg, Pawluk, Gawlowska, & Proudfit, 2014) – if IU represents a very circumscribed form of fear and anxiety variance, its predictive strength might be decreased if other motivational factors are, intentionally or not, activated in an experimental task.

Finally, it should also be noted that the remaining narrowband traits of interest, trait anxiety and anxiety sensitivity, were both components of interactions with the Stimulus dimension that significantly predicted avoidance, but plotted generalization slopes and fitted lines did not support an interpretation consistent with generalization. More specifically, the statistical interactions appeared to be driven by a shallower decline in responding across the stimuli more distal from the CS+ (GS2 to oCS-) for those higher on the traits, however, overlapping CIs for these portions of the gradient indicated this was likely not a replicable association and did not represent a generalizable result. It is possible that this is a genuine effect, but weak enough that statistical uncertainty remains.

Additional findings and Aim 1 summary.

In addition to the findings discussed above, there were a number of intriguing secondary findings that were either components of models containing multiple personality variables or exploratory analyses. The majority of these findings related to Extraversion, its aspects, and pathological extensions or manifestations of the trait. Perhaps most noteworthy is that Detachment was a significant predictor of increased behavioral fear generalization and avoidance generalization in a model that also contained Negative Affectivity and Disinhibition (all three are PID-5 traits). Detachment is a pathological personality trait capturing extreme disinterest, anhedonia, and social withdrawal that was included as a pathological correspondent to Extraversion, with lower

Extraversion corresponding to the lower-end of the Detachment distribution (which is observed in our sample as a correlation of $r = -.61$ between Detachment and Extraversion). It is reasonable to predict that Detachment would be related to worse discrimination on the task due to generally less engagement (i.e., those higher on Detachment are not motivated to expend energy to rigorously learn the task contingencies) and higher avoidance during danger due to lower reward motivation that could buffer against risk aversion. Behavioral generalization (i.e., risk rating) results appear to indicate that higher Detachment is associated with increased generalization from the GS2 to the GS1, which might be driven by poorer discrimination between these similar stimuli. However, we are cautious to extensively interpret this interaction due to 1) a lack of explanation for why poorer discrimination wouldn't be evident for the response slope from the GS2 to the GS3 and GS2) the CIs between the fitted lines for this interaction overlap, indicating that this result is potentially limited to our sample and not generalizable to the population. More interpretable is the interaction effect indicating Detachment is associated with generalized avoidance, which is evident as a shallower slope from GS3 to oCS- associated with higher Detachment. This appears to indicate that as the approach-avoidance conflict associated with each stimulus becomes clearer (i.e., less similar to the most ambiguous approach-avoidance conflict situation, the GS3), the tendency for those high on Detachment to not engage with reward becomes increasingly evident. Consistent with Detachment's conceptualization as a pathological trait, this lack of engagement in the context of the stimuli that more resemble the safety cue than the danger cue can be considered an index of increasingly maladaptive avoidance. That said, we cannot be sure that this effect is driven by the pathologically lower positive affect in

Detachment, as there is still a significant negative affect component to Detachment (Negative Affectivity and Detachment are highly correlated, $r = .76$). However, the fact that the Detachment interaction is significant in a model that also includes a significant Negative Affectivity interaction is notable, as this indicates that at least some of the negative affect-related variance in Detachment has been accounted for by the Negative Affectivity interaction.

The significant Enthusiasm-PB interaction predicting decreased behavioral fear generalization can be interpreted similarly to the Detachment interaction, as lower Enthusiasm-PB conceptually aligns with higher Detachment, and the interaction is defined by those lower on Enthusiasm-PB generalizing at the GS2 and GS1 levels, whereas those higher on the trait do not. In fact, it appears these interactions are graphically identical. Give that 1) Enthusiasm-PB is comprised of multiple facets that contribute to Detachment scores (e.g., Anhedonia, Intimacy Avoidance) and the resulting extremely high correlation between the two scales ($r = .95$ when aligning both scales so that higher scores = more extreme pathology) and 2) that again we see overlapping CIs in the fitted lines which limit confidence in this effect, it does not appear that further interaction of this particular interaction is warranted. However, the lack of a corresponding avoidance generalization effect for Enthusiasm-PB, along with the complete lack of Extraversion (and its aspects) effect on avoidance generalization, support the notion that normative levels of Extraversion and related dispositions are not inherently protective against maladaptive avoidance, and that the pathogenic capacity of Detachment to increase avoidance generalization is likely not related to its shared variance with Extraversion.

Although we did find avoidance generalization effects, lower Extraversion and Assertiveness-PB significantly predicted physiological fear generalization, with lower Extraversion associated with a modest degree of generalization limited to the stimuli more distal from the CS+ (GS2 and GS1), and Assertiveness-PB associated with a greater degree of generalization across the continuum. This is consistent with the conceptualization of Extraversion representing normative variation and Assertiveness-PB extending into the pathological tail of the Extraversion distribution, and that therefore the Assertiveness-PB is serving in the “full-spectrum” capacity that served as the theoretical rationale for using these composite variables. To enhance interpretation, it is worthwhile to reiterate what differentiates Extraversion from Assertiveness-PB: both contain the Assertiveness items from the BFAS, but Extraversion is also comprised of the Enthusiasm aspect, whereas Assertiveness-PB also includes Submissiveness, a pathological facet that is part of the PID-5. Thus, it can be reasonably inferred that the shared features between the two interactions (generalization at the level of the GS2 to the GS1) are related to their shared variance (i.e., the Assertiveness aspect), and that the addition of pathological Submission variance (conceptually opposing Assertiveness, $r = -.27$ in the current sample) is related to the stronger generalization effects that extend to the GS3 seen in the Assertiveness-PB interaction. A possibility is that introducing the pathological variance of Submissiveness enhances the Assertiveness trait so that it is more related to higher negative affect (which is consistent with correlations from this study, as negative correlations between Assertiveness-PB and the negative affect-related variables are uniformly stronger than those between Assertiveness alone and those same variables), which in turn appears more related to fear generalization from the CS+ to the

GS3. In that case, it is likely there are multiple determinants for the observed generalization effect: generalization for stimuli more resembling the CS+ is more related to high negative affect, generalization for stimuli less similar to the CS+ (or, alternatively, more similar to the CS-) is more related to low positive affect. A possible explanation for this latter effect is that those who have a higher threshold for activation of unnecessary defensive responding (i.e., lower “false alarm” rate during safety) are more assertive in their environments due to a lower rate of inhibitory signaling, and that this would be one of the determinants for higher scores on the Assertiveness scale. This is partially supported by evidence that those higher on Extraversion or related traits show overall decreased defensive reactivity compared with those lower on Extraversion (e.g., Corr, 2002), as well as by evidence of increased threat discrimination (as indexed by enhanced extinction) in those higher on Extraversion (S. L. Rauch et al., 2005) and decreases in acquired fear maintenance in those lower on the Activity facet of Extraversion (Pineles et al., 2009). That said, this does not explain why Extraversion is related to physiological generalization, but not Assertiveness alone, especially if we propose that Assertiveness is the aspect driving the observed generalization effect. This discrepancy could be neatly resolved if low Enthusiasm was strongly related to Submissiveness, but this is not the case ($r = -.05$), nor is it expected given how the composite scales were constructed (i.e., based on factor analytic techniques aimed towards optimizing the discriminant validity of each factor in regards to the other factors; DeYoung, Carey, Krueger, & Ross, 2016). Another possible, if tenuous, explanation is one focusing on an equifinal system in which low Assertiveness alone is not sufficient to elicit generalization effects, and that an additional source of variation related to increased

negative affect is needed, but that the particular form of generalization is differentially related to the source negative affect. In this case, it is possible that the low Enthusiasm which contributes to the lowest Extraversion scores is providing that negative affect (Enthusiasm is consistently anticorrelated with the negative affect traits to a greater degree than Assertiveness), but that, in contrast to Submissiveness, this source of negative affect variance is not sufficient to elicit the form of generalization most commonly linked to pathology (generalization from the CS+ to the GS3). This explanation is, of course, quite preliminary and would require substantial follow-up analyses and investigations to substantiate.

Overall, Aim 1 results confirm that we can predict fear and avoidance generalization using continuously modeled personality variables. The exact nature of these relations is complex, but a few overall trends can be extracted from the amassed results. First, in terms of normative personality, the Neuroticism and Conscientiousness-related effects all indicate that higher levels of these dispositions contribute to generalization in some fashion, but that the exact form of generalization differs, which also suggests that the underlying mechanisms also differ. This statement is also somewhat weakened by inconsistent findings across fear and avoidance generalization for these variables. Second, Extraversion-related effects suggest that higher levels are protective and buffer against fear, but not avoidance, generalization. Third, personality variables that capture pathological variance (both broadband and narrowband) were consistently linked with fear generalization, but narrowband measures were not associated with generalized avoidance, potentially indicating that the broader measures

contain content more associated with avoidance. Taken together, these results are intriguing, but require further investigation and replication.

Aim 2: Testing improved prediction of avoidance in APIC models

As stated, Aim 2 revolves around the question of whether we can further improve our avoidance prediction once we know someone's level of fear or threat estimation. The answer appears to this appears to be "yes", as we found generally consistent evidence that both broadband and narrowband traits can be added to predictive models to improve explained variance. Interpretation within this aim is limited given potential redundancy with interpretations for effects found for other aims.

A2.H1.

Our hypotheses regarding negative affect traits were partially supported: Neuroticism significantly improved model fit when added to both types of APIC models (physiological and behavioral Pavlovian indices), and Distress-PB significant improved model fit for the APIC model with startle, but not risk ratings. Unexpectedly, we also found that Volatility significantly improved model fit for both types of APIC models, and to a significantly greater extent than Neuroticism. This is partially reflecting the finding from Aim 1 that Volatility predicts avoidance and therefore including it in an APIC model overall improves the predictive power of the model, but also suggests that Volatility is adding uniquely predictive variance above and beyond the Pavlovian predictor. Further, Volatility is not significantly related to the Pavlovian responding, either overall or generalized. Therefore, it seems reasonable to conclude that Volatility is reflecting a separate source of variation that strongly relates to avoidance but is not related to Pavlovian responding or APIC. One possibility is that those higher on

Volatility are also those most likely to avoid as a form of defiance or refusal to participate in the “rules” of the video game. To some degree, this aligns with Volatility containing emotional lability and vulnerability variance, which is commonly associated with antagonistic traits and externalizing problems (e.g., DeYoung, Quilty, & Peterson, 2007; Wright & Simms, 2014). Even more speculative is that Volatility is tracking narcissistic qualities in a subset of participants (Volatility as measured on the BFAS has been previously linked to narcissistic personality dimensions, e.g., Oltmanns & Widiger, 2018) and that those participants do not emotionally identify with or care about the farmer getting to his crops. Either of these explanations might account for the improved predictive properties of the APIC models with Volatility without any corresponding relationship to Pavlovian responding, but at this time these explanations remain speculative.

A2.H2.

Although Conscientiousness itself did not improve APIC model fit, our hypothesis that Industriousness and Orderliness included in the same model would result in a significant improvement in model fit was supported when Pavlovian responding was operationalized with startle. Further, when examining the corresponding coefficients from the model with both aspects, there was a negative predictive relationship between Industriousness and avoidance ($\beta = -0.2579$) and a positive predictive relationship between Orderliness and avoidance ($\beta = 0.3464$). It appears that the expected suppression effect occurred and thus we can observe the expected relations between the two aspects and avoidance. This suggests that even above and beyond modeling APIC, there is added utility when including both Industriousness and Orderliness to predict avoidance. Put

another way, this model suggests that even if we know a person's level of fear, we can still refine our avoidance prediction if we know both their levels of Industriousness and Orderliness. Given the lack of any other Industriousness effect found in this dissertation, this finding suggests that Industriousness and its hypothesized enhancement of approach tendency is only apparent when accounting for multiple other sources of variation, and is therefore a subtle and potentially weaker effect. Finally, the fact that this effect was only evident for APIC models including startle complicates interpretation, and within the framework of the Aim 2 analytic technique (i.e., model comparisons using LRTs) suggests that Orderliness and Industriousness provide added predictive power only when accounting for physiological Pavlovian responding, or conversely, these aspects do not enhance prediction when a person's threat estimation (risk ratings) is known. Risk ratings represent a more volitional, cognitively-mediated form of Pavlovian responding (e.g., Beckers, Krypotos, Boddez, Effting, & Kindt, 2013; Boddez et al., 2013). Similarly, Conscientiousness, the superordinate trait for Orderliness and Industriousness, is frequently defined by and related to substituting volitional, organized responses for prepotent, automatic responses (e.g., Jensen-Campbell et al., 2002; McCrae & Löckenhoff, 2010). Therefore, it is possible that the information provided by risk ratings is partially redundant with that provided by the Conscientiousness aspects, and results in a lack of significant model improvement.

A2.H3.

Contrary to our hypotheses, adding Extraversion or its aspects to APIC models did not significantly improve prediction of avoidance. However, an additional finding for which we did not have specific predictions might help explain this lack of findings:

Assertiveness-PB significantly improved prediction for the APIC model that included risk ratings. This suggests that pathological, but not normative, forms of Extraversion improve APIC prediction. That said, we limit our interpretation at this point in the dissertation to avoid redundancy with interpretation of multiple Extraversion-related APIC findings in the next session.

A2.H4.

As with our prediction for broadband negative affect traits, our prediction regarding narrowband traits was partially supported. Both IU and trait fear significantly improved model fit for APIC models whether operationalizing Pavlovian responding with startle or risk ratings. As both of these variables were associated with increases in overall avoidance while also related to either overall or generalized Pavlovian responding, it is likely that they represent one of the most predictive variables in this endeavor and that their inclusion in APIC models improved model fit despite some overlap with Pavlovian variance. More interesting is that trait anxiety also significantly improved model fit for both APIC models, but anxiety sensitivity did not. A clear and parsimonious explanation for this discrepancy is not readily apparent. A more speculative possibility is that ASI, our measure of anxiety sensitivity, is along with the IUSF one of the more narrowly circumscribed narrowband measures used in this dissertation, but unlike the IUSF, the ASI does not contain items measuring avoidance or escape (Reiss et al., 1986). Therefore, it might be too specialized to improve avoidance prediction, but also not broad enough to capture variance that is distinct from the other predictors in the model.

Additional findings and Aim 2 summary.

In addition to the above discussed results, we found that APIC models containing all the aspects of interest (referred to as a “multi-aspect” model), as opposed to the aspects modeled separately or in pairs, resulted in significant model improvements. Further, these multi-aspect models yielded the largest increases in variance explained by the model compared with the base model whether the Pavlovian predictor was risk ratings or startle. This is likely partially due to these multi-aspect models containing more variables than the other models (5 more than trait/single aspect models, 4 more than the models containing both aspects, 3 more than the models containing all traits), which inevitably results in an increase in our R^2 metric, although the marginal R^2 used for the multilevel models in the current study is relatively robust to this issue (Nakagawa & Schielzeth, 2013). It is therefore also likely that part of this improvement reflects a genuine advantage of including multiple personality predictors in an APIC model to improve avoidance prediction. Further, this lends support to the notion that using aspects instead of higher-order traits allows more precise prediction of behavioral outcomes, and that absence of findings in prior studies might be related to only using trait-level variables (as documented in Lonsdorf & Merz, 2017).

Overall, the value of these results appears to primarily be as a “proof-of-concept”: the approach used in Aim 2 holds promise as a tool for answering preliminary questions related to personality variables and APIC. The exception to this conclusion is that we can start to see broad, but interpretable, patterns emerge from these results when considered as a whole, particularly as it relates to what types of personality variables (as opposed to which specific variables), most improve the predictive properties of APIC models. From that point of view, it appears that the relatively more narrowband variables (if considering

the aspects as representing narrowband variables relative to the traits) provide the greatest increase in prediction while still maintaining an acceptable level of model parsimony.

Aims 3 and 4: Moderators of APIC, APIC-G, and APIC-CS+

Due to their conceptual and statistical overlap, we discuss Aims 3 and 4, and the limited number of significant moderators identified, in the same section. These are perhaps the most difficult results to interpret due to APIC moderation being operationalized with a three-way interaction (Trait x Stimulus x Pavlovian variable) that statistically has notable limitations (discussed in a later subsection). Therefore, we consider these the most preliminary findings in this study, and thus those that are interpreted with the most caution. As with Aim 1, but to a greater extent, significant effects involving startle were largely absent.

A3.H1/A4.H1.

Contrary to our hypotheses, there was no evidence of significant APIC moderation by Neuroticism or Distress-PB. However, a significant moderation effect was evident for Negative Affectivity for APIC-G when operationalized with Pavlovian reactivity with a behavioral measure (risk ratings). This appears to partially support our assertion that the conversion of generalized Pavlovian responding to generalized avoidance when a genuine threat is absent is a pathological process, as the only negative affect measure we found to be related to APIC-G is specifically one that captures pathological variance; those purportedly containing normative variance (Neuroticism, its aspects, and Distress-PB) were not associated with APIC-G. On the surface, this appears to be the first evidence that broadly defined, pathological negative affect enhances the APIC-G process (i.e., is a risk factor for an increased maladaptive fear-avoidance

relation), and that is largely true based on a broad interpretation of this finding. However, more in-depth interpretation is needed to qualify this finding. First, individual analysis of each GS within the APIC-G analysis revealed that higher Negative Affectivity clearly facilitated the GS1 and GS2 Pavlovian-Instrumental association, which is consistent with our prediction. However, at the level of the GS3, the moderation effect was smaller in magnitude and appeared to be in the opposite direction, with *lower* Negative Affectivity facilitating the association. There appear to be multiple components that explain this unexpected result. First, the significant interaction term is clearly driven by those higher on Negative Affectivity, both in terms of statistical and graphed results – this can be seen in the simple slopes plot, in which the slopes for high, but not low, Negative Affectivity varies as a function of Stimulus. The high Negative Affectivity slope at the level of the GS3 is flatter than those for GS2 and GS1; the slopes for low Negative Affectivity remain the same across all stimuli. Closer inspection of this plot and the accompanying scatterplot reveal that the flatter high Negative Affectivity slope for GS3 corresponds to a cluster of participants who were predicted to avoid at a higher rate (e.g., 25% to 45%) despite relatively lower predicted risk ratings (e.g., the mean to 1 standardized unit below the mean). This means that a subset of participants with higher Negative Affectivity were avoiding at a relatively high rate despite their relatively low threat estimation, and potentially indicates that, for stimuli most similar to the danger cue, Negative Affectivity exerts its pathological influence through increased generalized avoidance without regard to threat estimation (i.e., the false alarm or “better to be safe than sorry” response). In other words, higher levels of Negative Affectivity might be associated with reflexively avoiding a stimulus that appears dangerous without considering the actual danger posed

by the cue, even if in previous encounters with this stimulus it has been learned that the stimulus is unlikely to be dangerous (i.e., overgeneralization). Related to this, Negative Affectivity was associated with both generalized risk ratings and avoidance in non-APIC models, and with both showing shallower declines from CS+ to GS3 associated with higher Negative Affectivity, potentially indicating that Negative Affectivity separately influences fear and avoidance generalization, but not the association between the two at the level of the GS3. That said, there is also a statistical interpretation that might help better explain these disparate results. The positive association between GS3 avoidance and risk ratings is relatively strong across all participants ($r = .33$), which is consistent with the GS3 eliciting considerably higher risk ratings and avoidance rates than the GS2; this is also where the “drop” is seen in most generalization gradients. Therefore, the experimental effect is perhaps too strong at this point in the stimulus dimension (i.e., the stimulus predictor accounts for the majority of the outcome variance), as the GS3 is the closest approximation to the CS+ and the majority of participants continue to rate it as non-benign and avoid it at a rate similar to their CS+ avoidance rate (as seen in the graphed generalization gradients presented as part of manipulation check analyses). If this is true, then there would be less between-subjects variance to predict at the level of the GS3. This possibility, combined with the subsample of higher Negative Affectivity participants that avoided at a substantially higher rate than expected based on their risk ratings, also suggests that there is a non-linear process occurring at the level of the GS3, and that additional parameters would be needed to adequately test the APIC relationship at the level of the GS3.

Also relevant to this interpretation is that the Negative Affectivity variable is fairly broad in terms of content, and includes items related to anxiety, depression, emotional lability, interpersonal insecurity, and other content with relevance to negative emotionality (Krueger et al., 2012). The structure of the variable and our analytic strategy precludes assignment of mechanistic significance to the fear and anxiety-related variance that is part of Negative Affectivity. It is therefore likely that the observed moderation effect is primarily a reflection of the higher sensitivity of Negative Affectivity to generally maladaptive processes, but that this also comes at the cost of specificity. That said, dismissing Negative Affectivity as too coarse to provide insight into more specific or complex processes is likely premature. Generalization processes (including, but not limited to, fear and avoidance generalization) are a hallmark of different types of internalizing psychopathology (e.g., Carver, 1998; Leung & Wong, 1998). For example, those with high levels of depressive symptoms might overgeneralize from an unsuccessful attempt to obtain a positive outcome, feel increased helplessness during incidentally similar situations, and subsequently not put in effort to obtain a reward that could have come at little to no cost (consistent with experimental accounts of learned helplessness and its generalization components; Mikulincer & Nizan, 1988; Seligman, 1972). Negative Affectivity is highly associated with internalizing pathology (Sleep, Hyatt, Lamkin, Maples-Keller, & Miller, 2018; Veith, Russell, & King, 2017). Therefore, it is possible that the moderation effect associated with Negativity Affectivity observed in this study is reflective of its covariation with the tendency towards conversion of generalized emotional reactions to maladaptive behavior, as opposed to an anxiety and fear-specific quality of the Negative Affectivity trait.

A3.H2/A4.H2.

Our hypotheses regarding Conscientiousness and its Orderliness aspect were largely supported, whereas we did not find any support for Industriousness as a moderator. In terms of the significant effects found, we exclusively interpret Orderliness over Conscientiousness, as 1) it was notably the only personality variable to moderate APIC when Pavlovian responding was operationalized either physiologically or behaviorally and 2) the moderating effect of Conscientiousness for the APIC model including risk ratings was close to identical to the Orderliness moderating effect for the same model, and therefore it is likely that the Orderliness aspect is driving the effect seen in the Conscientiousness results.

The moderating effect of Orderliness was most consistently seen at the levels of the GS1 and GS2, with lower levels of Orderliness facilitating APIC for these stimuli. In contrast, those higher on Orderliness show facilitated APIC for the GS3 (for startle only) and CS+. Further, the only instance of a personality trait *attenuating* APIC is seen for Orderliness, with those lower on the trait demonstrating a negative relationship between startle and avoidance. To briefly summarize, it appears that as Orderliness decreases there is an increase in the strength of APIC for the stimuli most resembling oCS-, and that as Orderliness increases there is an increase in the strength of APIC for the stimuli most resembling the CS+. One plausible explanation for this effect is that those lower on Orderliness might have also been the participants least likely to efficiently track the “rules” of the task and misidentify safety stimuli as having been previously paired with shock, believe that there were separate reinforcement parameters governing the CS+ and the different CS-s (e.g., CS+ reinforced 100%, CS- reinforced 10%), or to apply an

idiosyncratic rule to the task that was not consistent with the available information (e.g., mistakenly assuming that shock is dependent on if you receive a startle probe or what level of risk rating you provide on a previous trial). This then might lead to an overcorrection on instrumental trials in which those lower on Orderliness are still unsure about the signal value of the GSs and are more likely to avoid even if their emotional response or appraisal of threat does not strongly suggest avoidance is required.

That said, the above explanation does not help us understand why lower Orderliness is associated with weakened APIC. Perhaps a more parsimonious and comprehensive explanation is that the observed moderation effect represents higher Orderliness being associated with generally the same APIC relationship across all stimuli (i.e., low variation between simple slopes), and lower Orderliness associated with less consistent APIC across stimuli (i.e., high variation between simple slopes). This implies that those higher on Orderliness are generally applying a more rigid rule in their avoidance decisions that results in avoidance responses more commensurate with Pavlovian responding, whereas those lower on Orderliness are either over or underutilizing their Pavlovian responses in the decision-making process while also relying on idiosyncratic rules or tendencies to inform their decision. This explanation is compatible with the previously outlined explanation for why those lower on Orderliness show stronger APIC for the GS1 and GS2, and also potentially helps explain why we see diverging APIC moderation relations between low and high Orderliness for the CS+ when Pavlovian responding is operationalized with startle: those higher on Orderliness continue to apply their consistent rule regarding avoidance decisions, but those lower on Orderliness are using other information or motivations to govern their decision when

encountering the CS+. This naturally leads to the question of why the CS+, compared with the other stimuli evokes a markedly different APIC relationship in those lower on Orderliness. If we consider that 1) there is evidence that Orderliness, and Conscientiousness more generally, covaries with tendencies towards controlling the internal and external environment to increased preparedness for negative emotionality and/or reduce it in the moment (Bartley & Roesch, 2011; N. T. Carter et al., 2016; Jackson et al., 2010); 2) as a corollary, those lower on Orderliness are less likely to have this motivation and other, competing, motivations could dictate behavior when faced with an aversive stressors; and 3) the CS+ is consistently eliciting the strongest Pavlovian responding across participants. Taken together, we might conclude that those lower on Orderliness are not motivated to control or downregulate distress when it passes a certain threshold, and that other motivations are primary during these situations. One possible motivation relevant to the CS+ is the motivation to confront and master one's fear, which in turn leads to an approach response when afraid (Putwain, Symes, & Wilkinson, 2017; Rachman, 2004), and might explain the finding that as startle increased, those lower on Orderliness were less likely to avoid for CS+ trials.

The final unaddressed element regarding Orderliness' role as a moderator is the finding that higher Orderliness facilitates APIC (operationalized with startle) at the level of the GS3. This appears to be readily explainable using the previously articulated logic: those higher on Orderliness might use a consistent or rigid rule of making avoidance decisions that align with their fear response, so if fear has generalized from the CS+ to the GS3, it then leads to avoidance. Although this rule leads to an adaptive decision on CS+ trials (avoiding a real threat), it becomes maladaptive on GS3 trials (forgoing a

reward when there is no actual threat), and therefore provides some evidence that the adaptiveness of a trait is dependent on the situation, and appears to reflect theoretical and empirical accounts of how extremely high Conscientiousness can become maladaptive and lead to pathology associated with overly rigid cognitions and behaviors that interfere with functioning (N. T. Carter et al., 2016; Samuel & Widiger, 2011).

A3.H3/A4.H3.

Results largely supported our hypotheses regarding Extraversion, its aspects, and APIC-G: those lower on Extraversion demonstrated facilitated APIC (Pavlovian responding indexed by risk ratings only) for the GSs, but not the CS+. More specifically, as the stimulus became less similar to the CS+, lower Extraversion moderated APIC more strongly (as seen in the increasingly steeper slopes for lower Extraversion in the corresponding simple slopes plots). Assertiveness moderated APIC in an identical fashion, indicating it is likely the aspect within Extraversion that is driving the observed APIC moderations. This leads us to conclude that as level of Assertiveness (and, more generally, Extraversion) decreases, maladaptive threat estimation (i.e., predicting harm when there is a safety signal) is more likely to convert into an avoidance decision that, given the context, is also maladaptive due to the unnecessary loss of reward. This can also be interpreted as higher Assertiveness functioning as a protective factor against maladaptive threat estimation converting into maladaptive avoidance. It is important to emphasize that those higher on Assertiveness did not necessarily never estimate any threat related to the GSs, as we can see in the data there were indeed some participants higher on Assertiveness who rated the GSs as somewhat dangerous (i.e., some risk of shock) – the protective element of Assertiveness appears to be that these threat

estimations did not result in a corresponding avoidance decision. This finding and its relation only to the GSs also helps establish Assertiveness as an adaptive trait in this context, as higher Assertiveness was not associated with a weaker APIC relationship for the CS+, indicating Assertiveness did not influence adaptive avoidance and, perhaps, that those higher on Assertiveness were able to appropriately inhibit their increased approach motivation to avoid harm. This might also explain why Extraversion is found to be a protective factor against pathological anxiety (Bienvenu et al., 2004; Jylhä & Isometsä, 2006), as those who are protected against maladaptive avoidance through increased approach behavior would likely not see reduction in valued activities, which would continue to reinforce their adaptive approach decisions and provide opportunities to further reduce residual anxiety association with fundamentally safe activities. It should also be noted that the lack of an Extraversion or Assertiveness moderating effect for APIC at the level of the CS+ potentially reflects that the majority of the variation seen in these scales in this study is normative and largely does not reflect the maladaptive extremes of Extraversion that are associated with taking dangerous risks and externalizing pathology (D. Watson, Stanton, Khoo, Ellickson-Larew, & Stasik-O'Brien, 2019; D. Watson et al., 2015). That said, it is also possible that the reward used in this study was not sufficiently hedonically salient to elicit the maladaptive approach patterns that are associated with pathological extremes of Extraversion.

The more specific APIC-G analyses align with the APIC findings to a certain degree, as those lower on Extraversion continue to demonstrate stronger APIC (when Pavlovian responding is indexed by risk ratings) at the levels of the GS1 and GS2; however, the moderating effect for the GS3 is not reflected in the APIC-G model. This

discrepancy is difficult to explain without partially ascribing some of the difference to the statistical approach used and its drawbacks: the APIC-G model includes a reduced stimulus dimension (only the GSs and the ΔCS^-), which results in a random intercept that is less informative than the one in the full APIC model. This has two primary ramifications for the current endeavor: the APIC-G model is overall an inferior fit for the data than the APIC model (and therefore our parameters, including the conditional effects we examine for moderation analyses, are estimated with increased error) and we are no longer accounting for CS+ responding in our model. A possible consequence is that the individual intercept that is estimated for each participant no longer accounts for what is likely a participant's maximal level of responding (i.e., the CS+ response), and therefore the intercepts now might be underestimating those with higher CS+ responding and overestimating those with lower CS+ responding for the GS3, which would potentially reduce variability in the GS3 APIC response and obscure the moderation effect seen in the full APIC model if CS+. Put more simply, when lacking the information provided by the CS+ in a model, it is more difficult to precisely estimate the GS3, which affects the ability to detect relations among the experimental indices and individual difference variables.

An additional discrepancy between the APIC and APIC-G models is that instead of Assertiveness appearing to drive the Extraversion moderation effect, as seen in the APIC model, we instead see the other aspect, Enthusiasm, with a pattern of moderation similar to Extraversion. The exception to this pattern is that higher Enthusiasm facilitates APIC at the level of the GS3, which is unique to these analyses and inconsistent with all other APIC/APIC-G models incorporating Extraversion or its aspects. This could also be

due to the previously outlined statistical difference in APIC vs APIC-G models and its possible consequences. Another potential explanation that is more psychologically-oriented is that, similar to our interpretation for Negative Affectivity, those lower on Enthusiasm are predisposed to avoid (perhaps due to lower reward sensitivity) and when faced with a more ambiguous threat situation (such as the GS3) they default to avoidance without consideration of other signals or information (i.e., their estimation of risk).

It should also be noted that both pathological extensions of Extraversion, Assertiveness-PB and Enthusiasm-PB, were significant APIC-G moderators with lower levels of the traits having similar facilitatory effects at the levels of the GS1 and GS2, and higher Enthusiasm-PB continuing to show the somewhat perplexing facilitatory effect for GS3. This appears to support the overall finding that traits with pathological variance emerge as more related to the GSs/generalization and APIC-G. It is difficult to provide a substantive interpretation of these effects that is distinct from the previous interpretations of Extraversion-related effects, given the previously discussed statistical limitations with the APIC-G model and the discrepancies between APIC and APIC-G Extraversion findings, and we therefore limit our conclusion of these effects to suggesting that this provides additional, but not meaningfully distinct, evidence for pathologically low Extraversion-related functioning as a risk factor for maladaptive decisions that result in unnecessary loss of reward.

A3.H4.

Contrary to our hypotheses, none of the tested narrowband personality traits significantly moderated either APIC model. Considering that three narrowband traits (trait fear, IU, trait anxiety) significantly improved overall APIC model fit (which only

represents an increase in proportion of avoidance responding accounted for by the model, and does not specifically indicate a moderation effect) but did not function as significant moderators, we have preliminary evidence that the narrowband traits in this study are generally predictive of avoidance when taking into account the Pavlovian response, but do not moderate the APIC relationship.

Aims 3 and 4 summaries.

When considering the APIC and APIC-G results together, instead of the individual components of each separate analysis, a pattern begins to emerge. First, Extraversion and related variables are the most consistently linked to APIC/APIC-G. This is overall consistent with the notion that the PIG paradigm places participants in an approach-avoidance conflict and that to predict performance during approach-avoidance conflict we can use a combination of variables related to both avoidance and approach for optimal explanatory power (Pittig & Dehler, 2018; Pittig, Treanor, et al., 2018). In a sense, Extraversion and related variables are providing the approach-related variable that has largely been missing from prior studies of fear generalization and avoidance that have largely focused on negative affect variables. Second, Conscientiousness and Orderliness, which are also not primarily linked with or defined by negative affect, were also significant moderators of APIC overall, but not APIC-G. This further supports the assertion that traits that are not explicitly negative affect related provide important information related to approach-avoidance decisions. Also, although we did not find any significant APIC-CS+ moderators, we observed that Orderliness moderated the APIC effect at the level of the CS+ in the APIC model. Taken together, we have preliminary evidence that Extraversion and related variables are most associated with protection from

maladaptive fear-avoidance relations that result in a loss of reward (i.e., maladaptive avoidance), whereas Conscientiousness (through Orderliness) is associated with protection from maladaptive fear-avoidance relations whether the outcome is maladaptive avoidance or maladaptive approach. In contrast, pathological Negative Affectivity and Detachment facilitated maladaptive fear-avoidance relations. To a large degree, this outcome is consistent with theoretical and empirical accounts of these traits and their relation to psychopathology: both Extraversion and Conscientiousness are well-established as protective factors in relation to Neuroticism and other negative affect traits, which are found to be risk factors for psychopathology (Andersen & Bienvenu, 2011; Kotov et al., 2010; Stanton & Watson, 2014). Further, it is sensible that Extraversion protects from maladaptive avoidance, but not maladaptive approach, as those higher on Extraversion (and, likely more importantly, its lower-level aspects and facets more closely related to approach motivation and reward salience) would likely be those more likely to seek reward despite high cost. This leaves Conscientiousness (through Orderliness) as the most consistently protective trait in this investigation. However, as opposed to Extraversion and the “more is better” relationship observed, it appears Conscientiousness findings are supporting a Yerkes-Dodson “inverted-U” view of the trait (Yerkes & Dodson, 1908), in which the high and low extremes are both associated with maladaptive outcomes (in this study, high Orderliness related to maladaptive APIC for the GS3, and low Orderliness associated with maladaptive APIC for the CS+, GS2, and GS1), but moderate levels of the trait are most consistently adaptive. That said, this relationship has also been observed for Extraversion (with extremely high Extraversion-related traits associated with mania and externalizing behavior, extremely low

Extraversion-related traits associated with depression and internalizing; Watson, Stasik, Ellickson-Larew, & Stanton, 2015), and it is likely that the PIG task (with its weak hedonic reward) is not well-suited to eliciting the pathological processes of high Extraversion.

Overall Synthesis of Results

The goal of this section is to consider the broader pattern of results that have emerged from our analyses for the three overarching personality dimensions measured (Neuroticism/negative affect, Conscientiousness, Extraversion) and then synthesize this with the literature. We begin with a qualifying statement: this endeavor represents a first step towards rigorous investigation of personality as it pertains to fear and avoidance generalization. As will be discussed in following sections, there are numerous improvements in methodology and analytic strategy that need to be taken before strong conclusions can be made, especially given that many of the effects of interest are complex and difficult to detect. What follows is, all things considered, a preliminary and cautious synthesis of results.

Broadband and narrowband negative affect traits.

The current study provides the first evidence of continuously-measured negative affective traits being positively associated with fear and avoidance generalization phenomena. Further, the traits with a greater focus on psychopathological extremes (e.g., Negative Affectivity, Distress-PB, the narrowband variables) were more consistently linked with generalization. This supports the theoretical and empirical models of generalization that propose it as a mechanism underlying anxiety and trauma-related conditions (Dymond, Dunsmoor, et al., 2014; Lissek, 2012). There is also evidence that

the scope of the negative affect measure matters in terms of the strength and form of the association with generalization phenomena: for example, trait fear and IU, which are conceptually closely related to anxiety pathology, were consistently associated with fear generalization, whereas Negative Affectivity and Distress-PB were associated with both fear and avoidance generalization, and also helped explain increases in maladaptive APIC. This suggests that multiband approach used in the current study can yield results that would not be possible to obtain if investigators focus only broad or narrowband measures.

In terms of alignment with prior work, current results indicating negative affect traits are associated with fear generalization contradicts previous studies that do not find associations between Neuroticism or trait anxiety and fear generalization (Arnaudova, Krypotos, et al., 2017; Torrents-Rodas et al., 2013). As previously stated in our review of these studies earlier in this dissertation, these prior studies did not use appropriate sample sizes or analytic techniques for continuous personality traits, and we therefore suggest that the difference in methodology between these prior studies and the current study is likely a determinant in the difference in results. As stated convincingly by Lonsdorf and Merz (2017), prior work in the fear conditioning field has not been optimized for precise individuals difference work, and we therefore contend that the current study represents an improvement over prior studies and presents evidence for “real” negative affect effects that is stronger than the evidence provided in other studies that these are null effects.

That said, we recognize that the negative affect findings are mostly confirmatory and less novel than other findings from the current study. Although necessary to test and report, we do not think that the insights afforded by the current findings will considerably

contribute to new theory and questions related to fear, anxiety, and other negative affect individual differences as they pertain to generalization. Put another way, considerable evidence and prominent theories already suggested that those with greater degrees of negative affect and internalizing traits show greater levels of fear generalization and avoidance (Dunsmoor & Paz, 2015; Lissek & van Meurs, 2014; Mineka, 1979; Mineka & Zinbarg, 2006; Pittig, Treanor, et al., 2018), and the current study primarily serves as framework for future investigations that will expand this line of work, as opposed to providing findings that expand our understanding of the psychological phenomena. The exception to this is our results related to the Volatility aspect, which are novel (and potentially contradict well-established theories related to emotional lability and anger; e.g., Harmon-Jones, Peterson, Gable, & Harmon-Jones, 2008) and suggest a need for more work clarifying the role of negative emotional traits and states that are distinct from fear and anxiety, such as anger and frustration.

Conscientiousness and related variables.

The current study emphasizes the need to incorporate individual differences variables that are not typically included in fear conditioning studies into future studies. The findings regarding Conscientiousness, its aspects, and the related composite variables, are a large part of this assertion. The findings related to the Orderliness aspect and APIC in particular are particularly compelling in this regard: we found what appears to be a non-linear relationship between Orderliness level and adaptive fear-avoidance relations, with both the higher and lower ends of Orderliness related to an increased chance of a maladaptive outcome depending on the context (i.e., the presented stimulus). As has been emphasized throughout this dissertation, what has in the past been

statistically and, sometimes, conceptually treated as error in prior generalization studies is partially comprised of psychological meaningful variance, and we now have evidence that we can “recover” some of that meaningful variance. Even if an investigator is not primarily interested in Conscientiousness, Orderliness, or similar constructs, there is still an argument to be made that, due to their differential influence on generalization and APIC processes, these variables should be measured and statistically accounted for in future studies. To illustrate this, consider if a researcher is interested in studying the generalized fear-avoidance relationship in PTSD using a military veteran sample. Evidence suggests that veterans with PTSD can be classified into latent profiles with distinct symptom patterns, some of which greatly differ on levels of Conscientiousness (Contractor et al., 2016). One of the most notable distinctions between the identified latent profiles is those with the overall strongest PTSD symptoms are those relatively lower on Conscientiousness and also have stronger externalizing tendencies, and those with less severe PTSD symptoms and with primarily reexperiencing and avoiding symptoms were relatively higher on Conscientiousness. Based on past studies and current study’s results, and assuming higher Conscientiousness somewhat corresponds to higher Orderliness for some of the veterans (a reasonable assumption given that many of those who have military careers score highly on measures related to Orderliness, such as self-discipline; Bilgiç & Sümer, 2009), we might see similar generalization results but differing APIC results: participants with PTSD will demonstrate greater fear generalization, but perhaps those low on Conscientiousness will show enhanced APIC as the stimulus increases in similarity to the CS-, and those higher on Conscientiousness will exhibit a more inflexible APIC relationship across the stimulus continuum. If this

Conscientiousness effect does exist in the data but the trait is not measured, it functions as noise and complicates interpretations due to increase heterogeneity within a purportedly homogenous group. This issue is not resolved even if switching to a dimensional approach, as the strength of the relation between continuously modeled PTSD symptoms and APIC will be attenuated by a confounding effect (i.e., the classic “third-variable problem”; MacKinnon, Krull, & Lockwood, 2000). Although this is clearly an oversimplification and does not take into account that there is likely a more complex and dynamic relationship between PTSD symptom severity, Conscientiousness, and generalization variables than articulated here, the point remains that we can reduce statistical noise by measuring and modeling relevant personality variables.

The current results also add to the literature on Conscientiousness and its effect on behavior, both in conditioning and non-conditioning studies. In terms of conditioning studies, the prior study by Pineles et al. (2009) found divergent associations between Conscientiousness facets and Pavlovian fear responsivity. Although we did not directly test differential fear response as in Pineles et al. (2009), nor did we find Conscientiousness or related variables were associated with Pavlovian response, it is interesting that in both that study and our study that lower-level variables within Conscientiousness demonstrate a heterogeneous pattern association with behavioral criterion. In contrast to Pineles et al. (2009), the study by Martínez and colleagues (2012) found that overall Conscientiousness negatively correlated with discrimination fear learning, which is again not reflected in our findings. It is noteworthy that the paradigms used in the two cited studies did not contain an approach-avoidance component, which possibly modulated fear responding in a dynamic fashion (e.g., decreased after

confirmation one can avoid the shock) and lead to results discrepant from past studies. Also notable is that both the prior studies and this study operationalize the fear response differently, which contributes to the difficulty in interpreting results across studies that has been noted by Lonsdorf and Merz (2017). In terms of non-conditioning studies, the association between Orderliness and maladaptive outcomes is somewhat consistent with studies that find Conscientiousness is negatively correlated with effective decision-making performance, both while under explicit duress (e.g., Byrne, Silasi-Mansat, & Worthy, 2015) and when asked to transfer prior learning to a new context (e.g., Studer-Luethi, Jaeggi, Buschkuhl, & Perrig, 2012). This potentially indicates that our results are reflecting an aberration in the decision-making process that is activated for those with higher levels of Orderliness when completing challenging or complex tasks. That said, we are cautious regarding our comparisons between the current results and the reviewed literature due to methodological differences, most pertinently that the majority of the previous studies examine Conscientiousness only at the trait level, and we cannot assume trait-level Conscientiousness results will replicate at the aspect or other lower-order level.

Extraversion and related variables.

As with Conscientiousness, Extraversion and related variables appear to help clarify noise and convert it into viable signal. We found, as expected, that including dispositional variables more attuned to reward and approach than fear and avoidance helped predict task outcomes, and that the tested variables generally operated as protective factors. This was consistently seen for Extraversion, as higher levels were associated with decreased physiological fear generalization and APIC/partial APIC-G, and the effects of its aspects largely aligned with this pattern (with the notable exception

for Enthusiasm discussed previously). Further, the pathological inverse of high Extraversion, Detachment, was associated with generalization and increased APIC, further underlining the protective role of Extraversion. This leads to a fairly intuitive conclusion: a stronger disposition towards reward and approach buffers against maladaptive approach-avoidance outcomes. This empirically-supported conclusion also might facilitate an increase in depression-focused or relevant fear and avoidance conditioning research, which is sorely needed when considering the overlap of anxiety and depression (L. A. Clark & Watson, 1991; D. Watson, 2005) and that GAD and PTSD, two disorders frequently associated with comorbid major depression (Kaufman & Charney, 2000), are also two of the disorders most theoretically and empirically associated with overgeneralization (Dymond, Dunsmoor, et al., 2014; Kaczurkin et al., 2016; Lissek et al., 2014; R. A. Morey et al., 2015).

The only prior conditioning study to find a significant relationship between an Extraversion-related personality variable and a conditioning outcome reported that two facets aligning somewhat with Enthusiasm and Assertiveness were associated with enhanced fear responding (Pineles et al., 2007). The other two comparable prior studies did not find Extraversion to be associated with a fear conditioning variable (Martínez et al., 2012; Otto et al., 2007). This puts our results at odds with the literature. We contend that, as suggested in reviews on the subject (e.g., Pittig, Treanor, LeBeau, & Craske, 2018), the lack of a reward-related experimental manipulation in the prior studies did not create a situation in which Extraversion-related individual differences in performance would emerge, and therefore that direct comparison between our results and these prior studies is not entirely productive.

Across-personality trends.

A pair of patterns in the data emerged that were generally related to individual differences, but not to one personality variable in particular. The first of these is that the overwhelming majority of our significant results are related to an effect that is driven by or localized to a non-CS+ stimulus. This is of course by design in the PIG, which is intended to study phenomena that are not exclusive to the CS+, but it also is consistent with the “strong situation” perspective of threat (based on “trait by situation” or “person by situation” perspectives from the social personality literature (Endler, 1977; Meyer, Dalal, & Hermida, 2010; Mischel, 1999) and its relation to fear and anxiety (Lissek et al., 2006), which posits that individual differences in fear-related responses are less likely to emerge during unambiguously threatening situations due to the activation of a normative and adaptive fear-response that is relatively uniform across individuals. The CS+ is likely a strong situation for fear, especially in the context of the PIG, and therefore the condition in which the fewest fear-related individual differences emerge. In contrast, the GSs are “weaker situations”, as the signal value of a GS is more ambiguous. In these weaker situations, there are fewer objective determinants to fear responding, so subjective factors, such as personality dispositions, determine responding to a greater degree than in strong situations. This pattern has been well documented in previous conditioning work (for reviews, see Beckers, Krypotos, Boddez, Effting, & Kindt, 2013; Lissek, Pine, & Grillon, 2006; Lonsdorf & Merz, 2017), yet the vast majority of these studies involve patient-control differences. Patient-control differences are a valid example of how weaker situations lead to individual differences, of course, but are also constrained by the weaknesses related to using categorical diagnoses and relying on

measures of central tendencies for inferential analyses that were outlined in the introduction of this dissertation. The current study therefore represents an incremental step forward for improved trait by situation work in the conditioning field.

The second pattern of note to emerge was, with a few notable exceptions, our significant results all involved the behavioral Pavlovian response variable, risk ratings, either as a dependent variable or as a predictor in an APIC model. If viewing this as a testament to the positive qualities of risk ratings, as opposed to problematic aspects of using eyeblink startle measurement (which can be found in the limitations section that follows), we can then consider why a behavioral measure of Pavlovian responding was frequently related to the personality variables tested in this study. One explanation is that online have high reliability due to a limit number of response options (3 in our study) and there are fewer sources of error than other, more complex forms of measurement, such as psychophysiological assessment. Another is that the personality variables in this study are more sensitive to the construct operationalized by risk ratings, threat evaluation/expectancy, than startle. This is plausible, as potentiation of the startle reflex is a valence specific measure that is mediated by the amygdala and, in the context of our experimental manipulation, reflects fear responding (Grillon, Ameli, Woods, Merikangas, & Davis, 1991). This is more circumscribed (i.e., likely has fewer determinants) than volitional threat evaluation (as measured through behavioral ratings), which receives input from a number of different neural circuits and cognitive processes (e.g., Boddez et al., 2013; Drabant et al., 2011; Grupe & Nitschke, 2011, 2013; Mathews, Mackintosh, & Fulcher, 1997). We can also consider that some of the traits measured in this study, such as Conscientiousness, contain content that is centered around self-control and regulation

(DeYoung et al., 2007; Roberts, Lejuez, Krueger, Richards, & Hill, 2014), which might be more strongly related to a behavioral response such as risk ratings than startle, as participants could potentially downregulate a prepotent response and provide a risk rating that is modulated by a reasoning process, but would not be able to do the same with their startle response. Finally, we also consider that unlike startle, there is a “right” or “optimal” response for risk ratings (i.e., “no risk” for all stimuli other than the CS+), and that this 1) might elicit a competitive performance motivation from some participants who want to correctly predict shock and internalize this as their primary task for the PIG, which might relate to some of the traits tested in the study; and 2) risk ratings might be more sensitive to maladaptive responding in general due to a natural “threshold” for what constitutes a maladaptive response, whereas maladaptive potentiated startle is always defined relatively (i.e., based on response to another stimulus or group of stimuli). This quality of risk ratings also potentially leads to ceiling effects for the CS+; we did not observe this as a problem in our data but it remains a possibility that there is a subtle ceiling effect for a subsample of participants with similar personality profiles that could affect results in an unforeseen fashion.

Benefits of a multi-trait, multi-level approach.

As a final part of our synthesis, we note that both single-trait and multi-trait analyses, for both broader and narrower sets of dispositional variables, yielded significant and potentially meaningful results. Further, although there was some redundancy, different model and variable types generally yielded different results with different interpretations and implications. This represents a general improvement over the prior work in the field, which typically exclusively relies on single-trait models (for review,

see Lonsdorf & Merz, 2017). That said, due to both the analytic approach used and the overall heterogeneity in our results, it is not feasible at this time to determine if single-trait or multi-trait, broadband or narrowband, is a superior approach. We instead suggest that inclusion of models with multiple traits that represent different levels of personality structure should function as the default for future studies, with single-trait analyses conducted only with the backing of a strong scientific rationale. This type of approach is highly consistent with ongoing efforts to improve classification and measurement of psychopathology (Kotov et al., 2017), and the current results suggest it is a feasible and potentially fruitful approach for future conditioning work.

Limitations

As with any study, this one contains a number of limitations of different degrees that we highlight so that future studies can improve on and correct these limitations. Given the ambitious nature and multiple novel components of this study, there are also a fair number of limitations that merit discussion. We therefore subdivide limitations into three broad categories: those related to the PIG and experimental procedures, those related to the statistical analyses, and those related to the choice of sample and individual difference measurements.

Experimental limitations.

The PIG paradigm and experimental approach for this study had a number of limitations. One of the most notable limitations is that the design of the PIG, which contains interspersed Pavlovian and instrumental trials as opposed to the blocked design (Pavlovian block → instrumental block) used in prior studies (Talmi, Seymour, Dayan, & Dolan, 2008; Y. Xia et al., 2019), prevents testing of a direct causal relationship in which

Pavlovian responding affects subsequent instrumental avoidance. Because of this limitation, we test for covariation between Pavlovian and instrumental responses, instead of a direct causal relationship between the two. This also creates the possibility that instrumental responses and the outcome affected future Pavlovian responding. For example, if a participant continuously avoids GS3 due to its resemblance to the CS+ they will engage in fewer opportunities to learn the GS3 is not dangerous, which might in turn continue to maintain or even exacerbate fear towards the GS3 (similar to the cyclic nature of fear and avoidance proposed in two-factor theory, Mowrer, 1951). In contrast, participants who choose to approach at a higher rate on GS3 trials might quickly learn its safety value and show a reduction in Pavlovian responding that is above what might be expected based on the US reinforcement schedule. Although this is a viable possibility, we contend that 1) this instrumental-to-Pavlovian effect, if present, is somewhat mitigated by continued reinforcement of the CS+ during the generalization phase, which likely prevents total extinction; and 2) that the benefits of the interspersed trial structure (less predictability, prevention of extinction, possibility of dynamic reinforcement learning) outweigh the possibility of the described effect. Also related to this issue is that higher rates of avoidance corresponded to a lower total number of shocks received, and therefore participants did not all receive the same number of shocks, which potentially introduces error related to levels of sensitization to the shock US, a non-associative process that could interfere with conditioning processes (e.g., Çevik, 2014; Greenwald, Bradley, Cuthbert, & Lang, 1998). A potential solution to this would be to include total number of shocks received for each participant in our models as a level 2 variable (i.e., individual difference); however, this also introduces the possibility of criterion

contamination in models predicting avoidance, as number of shocks received will strongly covary with avoidance rate.

Another notable limitation is related to the reward component of the PIG, both in terms of the form of the reward and the static nature of the conditions required to obtain the reward and its association with the stimuli. As the original intent behind the design of the PIG was to elicit Pavlovian and instrumental avoidance generalization in the same task (van Meurs et al., 2014), the requirement for a reward was simply that it did not overpower or render the shock irrelevant. For the current investigation, however, this limits our ability to draw strong inferences regarding the effects of reward motivation and how they relate to candidate personality variables. For example, questions regarding if there is a certain level or threshold of reward intensity or likelihood that will buffer against maladaptive avoidance are not answerable with the current data. Further, a recent study has established that experimental tasks can elicit reward generalization gradients in participants (Andreatta & Pauli, 2019), and deficiencies in reward generalization have been linked to internalizing pathologies (Radell, Beck, Gilbertson, & Myers, 2017; Rouhani et al., 2018), suggesting that there are reward generalization processes that could be relevant to the current study (and potentially a determinant for some of the observed effects) but are not testable with the current data. Additionally, our conceptualization of the PIG paradigm as involving a risk-reward decision and potentially providing insight into decision-making processes is somewhat hampered by the constant nature of the award, as a static reward probability does not allow us to test hypotheses related to risk or loss aversion, which have great relevance to the personality traits tested in this study and psychopathology (e.g., Glimcher & Rustichini, 2004; Kishida, King-Casas, & Montague,

2010; Lauriola & Levin, 2001). It is also likely that the reward was not hedonically salient enough to probe for maladaptive approach behaviors that might be associated with Extraversion (D. Watson et al., 2019).

A further issue for the current study related to the PIG paradigm is that, as a generalization task, the PIG requires a relatively large number of different trial types (e.g., stimulus, Pavlovian vs instrumental) and permutations of trial type parameters (e.g., Pavlovian trials in which risk ratings are assessed or not), which results in a need to reduce the number of trials per type. Further, the upper limit of total number of trials is not only determined by concerns about participant performance declines due to boredom or fatigue, but that over time the startle reflex will habituate until the signal is not detectable (Blumenthal et al., 2005). This results in the PIG containing a large number of trial types, but a relatively small number of instances of each trial type. This had two consequences relevant to the current study. First, even one trial of missing startle data potentially created reliability issues for the associated trial type (Lieberman et al., 2017). This was somewhat mitigated by the robustness of MLMs regarding missing data (C. Krueger & Tian, 2004), but encouraged us to take a conservative approach to missing data (i.e., pair-wise removal). Second, it prevented us from using a three-level MLM to model the time component of the PIG (that is, Time/Trial would be nested within Stimulus Type, which would continue to be nested within Person), which would afford insight into the temporal dynamics underlying generalization and avoidance (and allow comparison to previous work on reinforcement learning as it relates to psychopathology, e.g., Mkrtchian, Aylward, Dayan, Roiser, & Robinson, 2017).

In terms of more general limitations that extend beyond the PIG parameters, a substantial number of datasets were removed prior to final analyses. Some data loss is to be expected in any empirical endeavor, and studies using psychophysiological measurement are perhaps more prone to data loss than many other forms of experimental testing (Cacioppo et al., 2007). That said, it is important to differentiate between those cases which could not be analyzed compared with those that potentially could be analyzed (i.e., our statistical models would fit the data), but raised concerns regarding the reliability or validity of their performance data. The majority of excluded cases were of the former type, but the latter type of exclusion is worth considering as a potential area for improvement and reassessment. Evidence suggests that performance-based exclusions can substantially alter interpretation of data and that rules for performance-based exclusions differ across research groups (Lonsdorf & Merz, 2017), and we want to emphasize that the exclusion criteria used in the current study could be further optimized.

Startle eyeblink assessed via EMG has been a fruitful technique for generalization work and for studies of emotion and psychopathology in general (Dymond, Dunsmoor, et al., 2014; Grillon & Baas, 2003). It also has documented reliability concerns (Larson, Ruffalo, Nietert, & Davidson, 2000), its elicitation requires the introduction of another aversive element to the experimental task (e.g., white noise burst, air puff), and individual-level startle data is typically not interpretable or useful. Although we contend that the benefits of startle (valence specificity, relatively easy and cheap to implement, well-mapped at the neural and behavioral levels) outweigh the drawbacks, it is fair to consider its limitations and how they might have affected the current study. Our primary concern is one of statistical power in relation to the ability to detect individual difference

effects in startle data. This concern is not unique to the current study; scientists have previously discussed the potential limitations of startle as it pertains to individual differences work (Grillon & Baas, 2002; Kaye, Bradford, & Curtin, 2016) and there are a number of published studies that did not find expected individual differences in fear processes as indexed by startle (for review, see Grillon & Baas, 2003). Although the current study has followed best practices for collection (Blumenthal et al., 2005) and quantification (raw startle magnitude used in our primary analyses, as recommended by Bradford, Starr, Shackman, & Curtin, 2015), we are cognizant of the possibility that type II errors occurred in relation to our startle data.

Statistical limitations.

Although the current study represents a continuation of a trend towards more sophisticated statistical methodology in the conditioning field, there are still limitations related to the statistical approach used. Some of these are more generally related to modeling generalization gradients, and some are more related to the individual difference aspect of the current study. Perhaps the most notable issue is related to the latter type of limitation: concerns about overfitting and interpretation prevented us from including a large number of personality variables, as well as personality variables from different measures, in the same model. A partial solution was the creation of the composite “PB” variables, but this was far from ideal. Even taking a relatively conservative approach, we contend that current statistical methodology was not ideal and that the issue of having a large number of models that are difficult to compare with each other strongly suggests that additional statistical techniques could have been used in conjunction with MLM to yield more parsimonious results. A combination of a linear and generalized multilevel

latent variables (e.g., multilevel structural equation modeling [SEM]) approach appears to be the ideal solution to this issue, especially given that these techniques allow for complex mediation and path analyses that can be modeled across different levels in the hierarchy (Hox, 2013; Mehta & Neale, 2005; Preacher, Zyphur, & Zhang, 2010; Rabe-Hesketh, Skrondal, & Pickles, 2004). An alternative to a latent variable approach would be to use regression techniques that minimize multicollinearity issues and assist with model selection (e.g., elastic net; Zou & Hastie, 2005) or to quantify the effect of leaving out individual or sets of individual difference variables and then manually decide on which predictors to include after obtaining an overview of the potential costs (e.g., the "Left Out Variable Error" approach; Mauro, 1990). However, we highly recommend the multilevel SEM approach be used in expansions of the current study, given the advantages afforded by the SEM framework and the substantial literature establishing latent factor models of psychopathology (e.g., Eaton et al., 2013; Rosellini & Brown, 2011).

Another set of limitations are related to the technique that we used to assess APIC-G. On a more technical level, the structure of the PIG resulted in an uneven number of individual risk rating measurements (18) compared with startle and avoidance measurements (36), which necessitated a reduced data structure in which the average of each trial type for each measurement was modeled, instead of individual trial data. This resulted in a poor fit for a random-slopes model, and therefore the APIC models do not contain random-slopes. The other limitation is related to our operationalization of what constituted APIC-G moderation (a three-way interaction that included the individual difference variable, the stimulus dimension limited to the Δ CS- and the GS, and the

Pavlovian variable) might be overly restrictive and have led to an increase in type II errors. This is because that, as constructed, there must be variation within the GSs for the interaction term to be significant. Thus, this *de facto* requires a significant APIC-G moderation to involve differentiated moderation effects at each level of the GS continuum, or at one level of the continuum in comparison with the other two GSs. Given that we found significant APIC-G moderators when using this more complex definition, it is likely that there exist additional personality moderators of APIC-G overall (i.e., across the GSs cumulatively, as opposed to individual moderating effects for each GS). One option for future analyses is to include an averaged GS term in the APIC-G model. However, there is concern about further reducing the stimulus dimension in MLMs to facilitate APIC-G analyses, due to the associated reduction in the number of available observations per level-1 cluster. This is also a possible explanation for the lack of significant APIC-CS+ moderators. A reasonable alternative might be to model each GS separately (which also provides additional information for fitting the random intercept and, if statistically prudent, a random slope) and include the average of the Pavlovian response to the Δ CS- as a level-2 predictor to function as a between-subjects covariate and control for baseline Pavlovian responding.

Another limitation is related to the hypothesis testing strategy used in the current study. The drawbacks of null hypothesis significance testing (NHST) and *p*-values are detailed extensively in the writings of many prominent statisticians and quantitative psychologists (e.g., Cohen, 1994; Wasserstein & Lazar, 2016), and those drawbacks apply here. It is also important to note that there are additional concerns about using NHST in the context of MLM (Gelman & Hill, 2006) above and beyond the standard

objections. These concerns are most focused on NHST for testing of individual coefficients, and include concerns about bias related to the maximum likelihood technique used to obtain parameter estimates, lack of consensus regarding appropriate degrees of freedom, and relatively arbitrary decision-making heuristics (e.g., not interpreting “nonsignificant” coefficients) reducing model utility. Compounding this issue, the interpretation of interaction terms in MLM is still an ongoing subject of debate and continued improvement among statisticians (Aguinis & Culpepper, 2015; McCoach, 2018). The current study used a technique that represented a compromise between the recommendations of the statistics community to move away from NHST, and the psychology community’s reliance on NHST and the simple slopes method of decomposing and probing complex interactions, as well as allowed for the specific aims of the study that related to testing precise interactions and limited the utility of model fit comparison methods of significance testing. Further analyses and investigations would benefit from more robust significance testing approaches (e.g., region of significance testing; Preacher, Curran, & Bauer, 2006) or moving to a Bayesian approach of modeling and evaluating data, which has seen strong support in the MLM community (Gelman & Hill, 2006; Hox, 2013) and is generally more robust and has fewer of the limitations seen in frequentist approaches to multilevel modeling (e.g., Gelman, Hill, & Yajima, 2012; Stegmueller, 2013).

Another potential limitation is related to the technique used to model the stimulus dimension. It is true that a categorically modeled and conceptualized stimulus dimension has substantial limitations, yet moving to a continuously-defined gradient might represent an overcorrection. It is possible that an *ordinal* dimension is the best fit for the goals of

generalization studies, as generalization researchers have specific hypotheses related to exact manifestations of the stimulus dimension (e.g., a close approximation to the CS+, but one that is still minimally perceptually distinguishable from the CS+, will elicit the greatest generalization of responding) but also conceptualize a meaningful order within the stimulus dimension (i.e., the GS3 elicits more generalization than the GS2, which elicits more generalization than the GS1, etc.). Techniques for representing ordinal variables in regression are established in the literature (Helwig, 2017) and represent an important next step for generalization research.

In the current study, we only used up to a second-order polynomial (i.e., quadratic function) in our models. This is consistent with the methods used and the results reported in prior generalization studies (see Dymond, Dunsmoor, Vervliet, Roche, & Hermans, 2014; Vanbrabant et al., 2015). It also might not be sufficient for precise analyses of individual differences in generalization, as a third-order polynomial (i.e., cubic function) might be the best fit for generalization data and therefore best suited to individual differences work. Consider the generalization gradients shown in this study and the vast majority of the cited references that represent empirical work on generalization: the gradients typically do not demonstrate a perfect curve from the danger cue to the safety cue. In most cases, there is a “bump” or elbow at some point in the continuum that represents the GS after which responding drops off precipitously. This suggests that incorporating cubic functions might be a worthwhile endeavor. Alternatively, inflection point techniques that identify the point in which the gradient begins its precipitous drop (e.g., Buss, Davis, Ram, & Coccia, 2018) could also be used and would likely provide

great utility without the increased difficulty in interpretation related to interpreting polynomial coefficients.

Sample and individual difference-related limitations

The most notable limitation related to the testing of individual differences is that we did not test models that included trait by trait¹⁸ interactions, which de facto means we did not compare if trait by trait interactions were stronger predictors of generalization or APIC than non-interacted personality variables (whether modeled individually or additively). Although our hypotheses did not hinge on trait by trait interactions and we were able to sufficiently accomplish our state aims without incorporating them in our models, this is still far from ideal if the overall goal is to move individual differences work in the conditioning field forward. Copious evidence supports the benefits of testing Big Five trait by trait interactions in the prediction of a variety of performance outcomes (e.g., McFatter, 1994; Silvia, Nusbaum, Berg, Martin, & O'Connor, 2009; Witt, 2002; Witt, Burke, Barrick, & Mount, 2002). More relevant to the current topic, it is a particularly valuable technique for predicting and characterizing psychopathological dimensions and their behavioral or biological correlates (L. A. Clark & Watson, 1991; Corr, 2002; Dinovo & Vasey, 2011; Jorm et al., 2000; L. C. Morey et al., 2002; Naragon-Gainey & Simms, 2017; Vasey et al., 2013, 2014). For example, consider that three trait-level dimensions highlighted in this work, Neuroticism, Conscientiousness, and Extraversion are found to interact such that higher levels of Conscientiousness and Extraversion can buffer against the development of clinical disorders in those with higher

¹⁸ “Trait by trait” reflects the common nomenclature in the literature; here we are referring to personality traits as a whole and not just the BFAS and PIF-5 trait-level variables.

levels of Neuroticism (e.g., Naragon-Gainey & Simms, 2017). This complex, but informative, interaction broadly aligns with the pattern of findings in the current study and suggests a need to incorporate this form of interaction testing in fear and avoidance generalization work. Modeling these types of interactions provides a method to approximate clinical disorders, even if not explicitly recruiting a clinical sample or oversampling for higher levels of psychopathological traits. For example, the measure of maladaptive personality used in the current study, the PID-5, was developed as an assessment tool to accompany the DSM-5 and to assist in measuring combinations of traits that contribute to personality disorder diagnosis (American Psychiatric Association, 2013; Krueger et al., 2012; Waugh et al., 2017). For example, antisocial personality disorder can be represented primarily with elevations of the Antagonism and Disinhibition traits and a selection of their component facets. These combinations of PID-5 dimensions can be statistically parameterized via interactive terms and tested as predictors of generalization phenomena in the current and future studies. This technique is not just limited to personality disorders, as the disorders that are frequently linked to generalization and conditioning abnormalities overall (anxiety disorders, PTSD) can be adequately defined with combinations of Big Five traits as well (Kotov et al., 2010).

Another limitation was our statistical approach to potential gender differences, which was to include gender as a covariate in every model. There are numerous gender effects noted in the literature underlying the current study, including fear and anxiety traits (McLean & Anderson, 2009), internalizing psychopathology (Nolen-Hoeksema, 2012), personality (Feingold, 1994), and conditioning processes (Rosenbaum et al., 2015). By statistically accounting for gender differences, we lose our ability to draw

conclusions about how gender influences the relationship between personality and fear and avoidance generalization, which could potentially improve our understanding of the gender discrepancy in diagnostic rates for anxiety disorders (McLean & Anderson, 2009). It is also possible that including gender as a covariate removed personality variation that is relevant to the current study but also is known to significantly covary with reported gender (e.g., select BFAS aspects; Weisberg, DeYoung, & Hirsh, 2011), suggesting an alternative approach might be required.

In terms of other non-tested variables that merit discussion, we did not include variables related to Agreeableness or Openness/Intellect, the last two traits of the Big Five, in our analyses. This was motivated by a lack of clear hypotheses or theoretical overlap with these traits and the experimental paradigm used, as well as concerns about an excessive number of tested predictors. That said, the Openness/Intellect trait, either overall or via one of its aspects, has been related to problem solving and working memory ability (DeYoung, Peterson, & Higgins, 2005; DeYoung, Shamosh, Green, Braver, & Gray, 2009), as well as psychological flexibility, emotion regulation, and mindfulness (Baer, Smith, Hopkins, Krietemeyer, & Toney, 2006; Gross, 2011; Kashdan & Rottenberg, 2010). All of these constructs have been evaluated as being negatively impacted by increased state or tonic anxiety and fear (Balderston et al., 2017; M. W. Eysenck, Derakshan, Santos, & Calvo, 2007; S. C. Hayes, Wilson, Gifford, Follette, & Strosahl, 1996; Kashdan, Barrios, Forsyth, & Steger, 2006), and some of this work has been conducted in the context of conditioning studies (R. M. Carter, Hofstötter, Tsuchiya, & Koch, 2003; Haaker et al., 2015; Jiang et al., 2018; Lissek et al., 2007). Taken together, it appears plausible that Openness/Intellect and associated subdimensions would

predict unique variation in one or more performance measures from the current study, and perhaps help clarify contradictory effects that could be related to deficits in cognitive abilities or flexibility.

Another notable omission is that we did not test a scale that can more directly assess reward approach and sensitivity, and thus could function as a positive affect equivalent to the narrowband variables used (which were related to the negative affect domain). An obvious selection for this would be the Behavioral Activation Scale (BAS) of the BIS/BAS framework (Carver & White, 1994), which works well in conjunction with Big Five measures to further flesh out reward-related variance with both normative and pathological relevance (S. L. Johnson, Turner, & Iwata, 2003; Smits & Boeck, 2006). It is also a key conceptual pillar of theories addressing approach-avoidance conflict (e.g., Corr, 2013).

Future Directions

One of the most important contributions of this dissertation is to the further development of experimental studies of fear and avoidance generalization. Both the significant results and the limitations of this study can helpfully inform that development, and the following represents our recommendations for the next steps in this line of work.

Optimizing the PIG.

The story of the fear conditioning field is, in part, a story of paradigmatic shift via improved experimental techniques (e.g., Beckers, Krypotos, Boddez, Effting, & Kindt, 2013; LeDoux, Moscarello, Sears, & Campese, 2017). Each investigation yields an opportunity to improve our methodology. In that spirit, the following is a list of potential changes or additions to the PIG that would further optimize the task for future work.

First, incorporation of a continuous measure of avoidance would provide more flexibility regarding analytic technique, as it would allow for movement away from use of the binomial distribution, yield increase response variation, and could be more naturally correlated with the ratio or ordinal-level Pavlovian responses measures. Examples of possibly relevant techniques can be found in the reward and effort-based decision making literature, and include speeded button press, repeated button press, or an effortful physical movement such as sustained grip intensity (e.g., Bonnelle et al., 2015; Hsu et al., 2015; Treadway, Bossaller, Shelton, & Zald, 2012; Treadway, Buckholtz, Schwartzman, Lambert, & Zald, 2009). In addition to conceptually aligning with the fact that avoidance motivation has many different behavioral manifestations and degrees, a continuous measure would also allow for parameterization avoidance in terms of both the decision to avoid (i.e., if the avoidance behavior is enacted at all) and the effort that is invested in avoidance (i.e., the effort invested in the behavior). Another advantage is that approach behavior can also be measured continuously using the same measure, depending on the specifics of the paradigm. As previously mentioned, manipulation of reward saliency or probability could also help diversify approach/avoidance responses, and when combined with continuously measured approach/avoidance data there are many possibilities for more nuanced and sophisticated analyses.

Second, we strongly argue for continued use of modern statistical techniques that are better suited for modeling individual variation than ANOVA techniques. This suggests that our experimental paradigms should be designed with the strengths and limitations of these statistical techniques as guiding principles. Accordingly, if using the MLM framework, we recommend that each level-1 predictor (i.e., within-subjects

measurement) have the same number of observations (e.g., same number of risk rating and startle measurements for each trial type). Ideally, each permutation of each trial type would also have the same number of observations (e.g., same number of risk rating and startle trials for each CS+, CS-, and GS), but this might not be feasible. Adhering to these recommendations would increase homogeneity across different models that include different Pavlovian or other level-1 predictors, which is crucial to ensuring the same random-effects structure is used for all models. On a similar note, reducing the number of distinct stimuli (e.g., moving from 3 GS classes to 2) while simultaneously increasing the number of trials per stimulus type would also likely result in superior model fit and cohesion across models. Further, this would likely avoid the situation in which certain stimuli are associated with low multivariate variance and therefore contribute a small amount of information to the model (as can be seen to an extent with the GS1 and oCS- in this study). An added benefit of reducing the number of distinct stimuli is that it would render the PIG more suitable for advanced computational modeling techniques used to assess temporal dynamics (Gershman & Daw, 2017; Mkrtchian et al., 2017) and allow testing of how these dynamics influence on task performance and provide a link to the influential literature that formally parameterizes the dynamics of Pavlovian and instrumental conditioning (Rescorla & Solomon, 1967; Rescorla & Wagner, 1972).

Another possible area for improvement is to consider the parameters of the US that are employed in the fear conditioning process and how this relates to potential strong or weak situations (Lissek et al., 2006). For example, if continuing to use shock as the US, establishing a relatively “low” and “high” level of shock for each participant and administering both using different reinforcement rates might allow for more precise

elicitation of fear and anxiety related individual differences as they relate to PIG outcome variables. Preliminary evidence indicates that this is an effective technique for capturing obsessive-compulsive individual differences in avoidance rates during a modified PIG paradigm (Hunt, Degeneffe, Bixby, Fleig, & Lissek, 2018). Shock is also not the only viable US, of course, and USs that are more aligned with specific traits of interest might help calibrate the PIG task so that it is optimally parameterized to elicit and detect individual differences effects. For example, social USs (e.g., negative faces and comments; Lissek et al., 2008) to probe effects related to Extraversion/Detachment, might be a fruitful next step to answer questions related to psychopathology with elevated fear to social stimulus (e.g., social anxiety disorder).

Finally, we note that we make an inherent assumption that participants are largely similar regarding individual motivations and learned rules during completion of the PIG, whereas the reality could be that there is a meaningful quantity of idiosyncratic variance that we are neglecting to measure. For example, some participants who have recently completed or are currently taking a psychology course might be primed to think about the role deception has played in famous, yet controversial, research studies that also use shock administration (Milgram & Gudehus, 1978), and assume that the PIG will involve some form of creative deception related to their avoidance decision (e.g., their performance affects another student completing a similar task). This, in turn, might bias their avoidance decisions and might also be related to some of the personality traits we test (e.g., Neuroticism), potentially introducing an unmeasured confound. Performance-based exclusions (e.g., exclude anyone rating the oCS- as riskier than the CS+) help with this to some degree, but are likely an overcorrection and contribute to an ongoing issue

related to inconsistently applied performance-based exclusion criteria across fear conditioning studies (Lonsdorf & Merz, 2017). One straightforward option to help resolve this issue is to simply ask participants open-ended questions regarding their motivations, beliefs, and learned rules in relation to the task. This could occur during the task if titrated appropriately, or via a more systematic interview that follows the task completion. Concerns about open-ended qualitative data are alleviated by the copious number of sophisticated techniques that exist to satisfactorily analyze these data (Silverman, 2016; Weston et al., 2001).

Incorporating additional measures.

In addition to optimization of the conditioning techniques and paradigm used in the current study, future work could benefit from additional measurement techniques to help clarify effects and answer more complex scientific hypotheses. One logical next step is to incorporate additional neurobiological measures, specifically functional neuroimaging. Functional magnetic resonance imaging (fMRI) and electroencephalography (EEG) are the two most common functional neuroimaging techniques used in the conditioning field (Büchel & Dolan, 2000; Lonsdorf et al., 2017), and both would provide valuable insights to future studies in this area. From an experimental point of view, the temporal resolution afforded by EEG is particularly attractive, as it could provide increased understanding of the rapid perceptual, attentional, and decision-making processes and that underlie approach or avoidance decisions and the feedback mechanisms that maintain or alter these decisions (e.g., Kessels, Ruiter, Wouters, & Jansma, 2014; Ratcliff, Philiastides, & Sajda, 2009). That said, fMRI remains increasingly popular and advantageous for fear and anxiety work (Etkin & Wager, 2007;

Fox & Shackman, 2017), as well as personality science (DeYoung et al., 2010), and is perhaps best suited to measure the fear and reward mechanisms that underlie performance on the PIG paradigm, given that fMRI's spatial resolution is conducive to accurate measurement of the subcortical circuits that are implicated in normative and pathological fear and anxiety (LeDoux, 2000). Technical differences aside, either functional neuroimaging technique would also yield more information regarding reward processes during fear and avoidance generalization, which is not possible when only using fear-potentiated startle. It should also be noted that if neuroimaging techniques are not available that there are other options for assessing biological correlates of reward that do not require expensive equipment and can be used in conjunction with the methodology described in the current study. Options include measurement of the postauricular reflex, a vestigial, evolutionarily conserved reflex in humans that pulls the ears back which no longer serves a behavioral function but is potentiated in the presence of pleasant or appetitive stimuli (Benning, 2011; Benning, Patrick, & Lang, 2004; Sandt, Sloan, & Johnson, 2009). Like the eyeblink startle reflex, the postauricular reflex is activated by quick acoustic probes, therefore making it an ideal measure to pair with fear-potentiated startle if interested in both fear and reward responding. Another option is to measure spontaneous blink rate, which has been used as an index of striatal dopamine-related activity (Groman et al., 2014; Peckham & Johnson, 2015) and could provide a tonic measure of reward responsiveness. Both techniques could be applied in the same study as well.

Another possible area of improvement is in the techniques used to assess psychopathology. Future studies that seek to compare their results to the large and

established work on DSM disorders might seek to include methods from that literature, such as structured or semi-structured clinical interviews (First, Spitzer, Gibbon, & Williams, 2002). These techniques are valuable even if adhering to a dimensional model of psychopathology, as they allow for more direct comparison to prior work and, if desired, the opportunity to demonstrate the empirical merit and advantages of a dimensional approach through this direct comparison. Additionally, there are advantages in using a specialized interview for specific symptoms, such as those associated with PTSD (e.g., Clinician Administered PTSD Scale; Blake et al., 1995), due to their added reliability and validity over self-report measures and ongoing work that uses interview-derived indices to identify dimensional models of PTSD that likely better represent the underlying pathology and relations amongst symptoms than the traditional DSM category (e.g., Armour et al., 2015; Miller et al., 2012). These types of interviews also frequently include questions that yield information about idiographic forms of fear and avoidance generalization (e.g., distress and avoidance in relation to boats, water, and any dizzying motion after nearly drowning following a boat crash), which could be used as external validity criterion to determine if experimental indices of fear and avoidance generalization align with personal, individualized experiences of the phenomena.

Implications for Clinical Science and Intervention

A primary proximal goal of the current research is to provide tools and knowledge that facilitate the next step in understanding generalization phenomena and their mechanistic relationship with normative and psychopathological dimensions of personality, particularly those related to anxiety, fear, and trauma. Our hope is that this work directly informs studies of intervention and change mechanisms, which in turn

inform formal treatment trials. Thus, the ultimate goal of this research is to facilitate translational efforts that eventually lead to improvement in the treatment of distress and impairment related to maladaptive fear and avoidance. We conclude this dissertation with a discussion of how the current research could ultimately inform clinical interventions.

The development, implementation, and optimization of exposure therapy, a family of interventions designed to reduce fear and avoidance through repeated exposure to feared stimuli in the absence of negative outcomes (e.g., Abramowitz, Deacon, & Whiteside, 2019), is one of the great success stories in the scientific tradition of clinical psychology. It is a story of ongoing and sustained effort to translate empirically-based animal models to human models (Jones, 1924; Pavlov, 1927; J. B. Watson & Rayner, 1920; Wolpe & Lang, 1964), which are then leveraged into an effective family of treatments for a range of psychopathology, most notably the anxiety and trauma-related disorders. In short, it is a sign that the “bench to bedside”¹⁹ model of translational science works (Starke, Fineberg, & Stein, 2019). It also continues to be a field of intervention in which basic psychological science work, especially those using conditioning techniques, can continue to contribute to meaningful improvements (e.g., Craske, Hermans, & Vervliet, 2018; Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014), but also one that requires some catch-up: the conditioning community largely ignoring avoidance phenomena for decades, and thus there was a lack of work that could be translated for clinical usage (LeDoux et al., 2017). Our study is part of a current wave of research to correct this imbalance (Pittig, Treanor, et al., 2018), and our results suggest multiple possibilities for improved intervention. That said, we acknowledge that considerable

¹⁹ In the context of clinical psychology, “science to session” is perhaps a more apt term

additional work that replicates and expands on our results is required before this work can provide a meaningful impact on the applied side of the scientific process.

First, we consider the construction of exposure hierarchies, in which situations that the client is avoiding are ordered from most to least distressing, is an important part of many exposure protocols. The organization of these hierarchies is typically not determined by anything more systematic than a self-reported rating of distress for each type of situation. It should also be noted that exposure hierarchies can be conceptualized as form of generalization gradient, as after client and therapist collaborate to generate situations, they are then ranked and re-rank according to level of distress, and then fine-tuned by increasing or decreasing the associated distress with a situation by slightly modifying the parameters of the situation (e.g., going to the mall alone or with a trusted friend). If we return to our example of the child and the Rottweiler, we can imagine that a therapist working with him would create an exposure hierarchy that contains situations such as “walking by a house with a dog at the window”, “watching dogs play in a dog park from outside the enclosure”, “being in the same room as a dog”, “petting a dog”, and so forth (Choy et al., 2007). The distress associated with each of these situations could be perhaps modified by specifying if the child is with their parent, or if the dog in question is dissimilar to the Rottweiler (e.g., the poodle) or more similar to the Rottweiler (e.g., the Doberman). Again, the key (and only) variable employed in this technique is subjective distress. Based on the current results, we might propose that a multidimensional system that is more sensitive to the personality traits and motivations that inform avoidance be used when indicated (e.g., for more complex clinical presentations). This system could take into account subjective distress, but for every avoided situation could also assess the

client's threat evaluation (e.g., estimating the chance that the feared consequence could occur in numerical units – a technique that is already used in CBT protocols; e.g., Beck, Davis, & Freeman, 2015), potential reward or valued outcomes related to the situation and the expected chance of the positive outcome occurring, and their belief of what “the right” way to handle the situation or what rules one is required to follow in this situation. These questions, in conjunction with formally (e.g., objective personality testing) or less formally (e.g., therapist-client interaction) obtained knowledge regarding the client's personality could help to anticipate and correct roadblocks in the exposure process before they occur. For an example, based on the finding of Extraversion protecting against APIC for the GSs most similar to the CS-, but not the CS+, we can imagine that the child from the ongoing example has been observed/measured to have a moderate level of Extraversion relevant to the population, and is indicating a large amount of distress and expected threat for the situations that are relatively dissimilar to the original harmful encounter with the Rottweiler, such as walking past a house with a dog that stays by the window, but also has listed a valued component of this activity, such as taking that route to arrive at a friend's house. The child has placed this activity relatively high on his hierarchy (i.e., will not be addressed until later in therapy due to the perceived challenge of the situation). Based on this combination of variables, we might conclude that although the child 1) reports strong distress associated with this situation and 2) it is quite dissimilar to the original incident itself, there is also a valued component involved, and that this situation should be moved down the hierarchy (e.g., completed earlier in treatment) because the child does not have a level of Extraversion that confers risk for increased avoidance. In other words, increased knowledge of the “trait by situation”

variables (e.g., Tett & Guterma, 2000) involved allows the therapist to more precisely assist in constructing the exposure hierarchy, which could ultimately lead to a decrease in dropout or poor exposure treatment outcomes. This is of course largely speculative and requires much more empirical work to substantiate, but the possibility of this type of multidimensional system for exposure being implemented is an on intriguing one. It should also be noted that many of the described concepts are familiar to CBT practitioners, but not formally applied to exposure work, which for many clients is the most potent mechanism of positive change (e.g., Boettcher & Barlow, 2018), and that a more rigorous exposure hierarchy construction that adds additional parameters might be particularly beneficial for those who do not see sufficient initial response to treatment.

Finally, we move outside the framework of exposure therapy and consider other interventions and psychotherapeutic approaches. Fundamentally, the current study represents an effort to 1) link a broad set of individual difference variables with clinically meaningful processes, even those individual differences not typically studied in the fear conditioning and behavioral exposure literature (e.g., negative affect traits); and 2) infer how different combinations of traits and processes could be considered adaptive or maladaptive based on the context (with context operationalized in the current study through the different experimental phases and trial types). This tracks with a goal of relatively newer psychotherapies to more thoroughly incorporate context and personal values into case formulation, such as Acceptance and Commitment Therapy (ACT; S. C. Hayes, 2004; S. C. Hayes, Strosahl, & Wilson, 2011), which centers on a “person as context” approach. The general framework of the current study aligns with this approach, as personality traits can be viewed as conferring information about personal values (e.g.,

someone higher on Extraversion likely values extensive social interaction to some degree) and the experimental conditions create variation in context and “adaptiveness” of an outcome.

In terms of results from the current student and their relation to contextual therapies, it is notable that one of our sets of findings was related to Orderliness, which at its pathological extreme is anticorrelated with psychological flexibility (Hewitt & Flett, 2007). Psychological flexibility is an individual difference that is posited as a higher-level determinant of health that protects against psychopathology and is proposed as a dimension that cuts across diagnoses and treatment disciplines (e.g., S. C. Hayes, 2002; Kashdan & Rottenberg, 2010; Latzman & Masuda, 2013). ACT and similar therapies purportedly confer therapeutic benefits through increased psychological flexibility, as opposed to the symptom reduction model that typically underlies exposure and CBT-based techniques. A highly speculative, but intriguing, idea based on the current results is that employing ACT or a similar approach to improve psychological flexibility in those extremely high on Orderliness might subsequently attenuate the covariation of problematic fear and avoidance in the client’s life. Clearly, this is conjecture at this point and would require extensive study to substantiate, yet we propose there is value in “thinking outside the box” and invoking a more creative mindset regarding experimental conditioning and individual differences work, and that innovation does not need to be limited to exposure techniques.

The increased push towards personalized medicine in the clinical psychology and psychiatry fields (e.g., Ozomaro, Wahlestedt, & Nemeroff, 2013; Schneider, Arch, & Wolitzky-Taylor, 2015) likely requires improved classification and delineation of individual variation in multiple domains to be successful – human complexity practically demands it. This, in turn, requires that experimental psychopathologists continue to investigate individual difference dimensions that cut across different diagnoses and meaningfully relate to maladaptive experimental outcomes that have relevance outside of the laboratory (Conway et al., 2019; Ruggero, 2018). Taken together, we are optimistic that the current study, and more generally the methodological and conceptual approach used herein, represents one such investigation that is vital to make the push towards personalized intervention into a reality. A further hope for this type of work is that it does not just inform improvements in exposure therapy and treatments for fear- and anxiety-based pathologies, but also contributes to the development of a wide range of effective personalized treatments that both acknowledge the complexity of psychopathology as it exists within the framework of human personality and leverages it into more effective treatment.

References

- Abramowitz, J. S. (1996). Variants of exposure and response prevention in the treatment of obsessive-compulsive disorder: a meta-analysis. *Behavior Therapy*, 27, 583–600.
- Abramowitz, J. S., Deacon, B. J., & Whiteside, S. P. H. (2019). *Exposure Therapy for Anxiety, Second Edition: Principles and Practice*. Guilford Publications.
- Acheson, D. T., Forsyth, J. P., Prenoveau, J. M., & Bouton, M. E. (2007). Interoceptive fear conditioning as a learning model of panic disorder: an experimental evaluation using 20% CO₂-enriched air in a non-clinical sample. *Behaviour Research and Therapy*, 45, 2280–2294. <https://doi.org/10.1016/j.brat.2007.04.008>
- Aczel, B., Bago, B., Szollosi, A., Foldes, A., & Lukacs, B. (2015). Measuring individual differences in decision biases: methodological considerations. *Personality and Social Psychology*, 1770. <https://doi.org/10.3389/fpsyg.2015.01770>
- Adelman, H. M., & Maatsch, J. L. (1956). Learning and extinction based upon frustration, food reward, and exploratory tendency. *Journal of Experimental Psychology*, 52, 311.
- Admon, R., Milad, M. R., & Hendler, T. (2013). A causal model of post-traumatic stress disorder: disentangling predisposed from acquired neural abnormalities. *Trends in Cognitive Sciences*, 17, 337–347. <https://doi.org/10.1016/j.tics.2013.05.005>
- Agresti, A., & Kateri, M. (2011). *Categorical data analysis*. Springer.
- Aguinis, H., & Culpepper, S. A. (2015). An expanded decision-making procedure for examining cross-level interaction effects with multilevel modeling. *Organizational Research Methods*, 18, 155–176. <https://doi.org/10.1177/1094428114563618>
- Aguinis, H., Gottfredson, R. K., & Culpepper, S. A. (2013). Best-practice recommendations for estimating cross-level interaction effects using multilevel modeling. *Journal of Management*, 39, 1490–1528. <https://doi.org/10.1177/0149206313478188>
- Ahmed, O., & Lovibond, P. F. (2018). Rule-based processes in generalisation and peak shift in human fear conditioning. *Quarterly Journal of Experimental Psychology*, 1747021818766461. <https://doi.org/10.1177/1747021818766461>
- Ahrens, L. M., Pauli, P., Reif, A., Mühlberger, A., Langs, G., Aalderink, T., & Wieser, M. J. (2016). Fear conditioning and stimulus generalization in patients with social anxiety disorder. *Journal of Anxiety Disorders*, 44, 36–46. <https://doi.org/10.1016/j.janxdis.2016.10.003>
- Aiken, L. S., West, S. G., & Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions*. Sage.
- Alarcón, D. E., Bonardi, C., & Delamater, A. R. (2017). Associative mechanisms involved in specific Pavlovian-to-instrumental transfer (PIT) in human learning tasks. *The Quarterly Journal of Experimental Psychology*, 0, 1–55. <https://doi.org/10.1080/17470218.2017.1342671>
- Allan, N. P., Cooper, D., Oglesby, M. E., Short, N. A., Saulnier, K. G., & Schmidt, N. B. (2017). Lower-order anxiety sensitivity and intolerance of uncertainty dimensions operate as specific vulnerabilities for social anxiety and depression within a

- hierarchical model. *Journal of Anxiety Disorders*.
<https://doi.org/10.1016/j.janxdis.2017.08.002>
- Allen, M. T., Handy, J. D., Miller, D. P., & Servatius, R. J. (2019). Avoidance learning and classical eyeblink conditioning as model systems to explore a learning diathesis model of PTSD. *Neuroscience & Biobehavioral Reviews*, *100*, 370–386.
<https://doi.org/10.1016/j.neubiorev.2019.03.003>
- Altman, D. G., & Bland, J. M. (1995). Statistics notes: absence of evidence is not evidence of absence. *BMJ*, *311*, 485.
- Altman, D. G., & Royston, P. (2006). The cost of dichotomising continuous variables. *BMJ*, *332*, 1080. <https://doi.org/10.1136/bmj.332.7549.1080>
- American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (Fifth Edition). Arlington, Va.: American Psychiatric Association.
 Retrieved from
<http://psychiatryonline.org/doi/book/10.1176/appi.books.9780890425596>
- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders*. (4. ed., text rev., 5th printing.). Washington DC: American Psychiatric Association.
- Andersen, A. M., & Bienvenu, O. J. (2011). Personality and psychopathology. *International Review of Psychiatry*, *23*, 234–247.
- Anderson, J. L., Sellbom, M., Bagby, R. M., Quilty, L. C., Veltri, C. O. C., Markon, K. E., & Krueger, R. F. (2013). On the convergence between PSY-5 domains and PID-5 domains and facets: implications for assessment of DSM-5 personality traits. *Assessment*, *20*, 286–294. <https://doi.org/10.1177/1073191112471141>
- Andreatta, M., & Pauli, P. (2019). Generalization of appetitive conditioned responses. *Psychophysiology*, e13397. <https://doi.org/10.1111/psyp.13397>
- Arch, J. J., & Craske, M. G. (2009). First-line treatment: a critical appraisal of cognitive behavioral therapy developments and alternatives. *Psychiatric Clinics of North America*, *32*, 525–547. <http://dx.doi.org/10.1016/j.psc.2009.05.001>
- Arkowitz, H., & Westra, H. A. (2004). Integrating motivational interviewing and cognitive behavioral therapy in the treatment of depression and anxiety. *Journal of Cognitive Psychotherapy*, *18*, 337.
- Armour, C., Tsai, J., Durham, T. A., Charak, R., Biehn, T. L., Elhai, J. D., & Pietrzak, R. H. (2015). Dimensional structure of DSM-5 posttraumatic stress symptoms: support for a hybrid anhedonia and externalizing behaviors model. *Journal of Psychiatric Research*, *61*, 106–113.
<https://doi.org/10.1016/j.jpsychires.2014.10.012>
- Armstrong, T., & Olatunji, B. O. (2017). Pavlovian disgust conditioning as a model for contamination-based OCD: evidence from an analogue study. *Behaviour Research and Therapy*. <https://doi.org/10.1016/j.brat.2017.03.009>
- Arnaudova, I., Kindt, M., Fanselow, M., & Beckers, T. (2017). Pathways towards the proliferation of avoidance in anxiety and implications for treatment. *Behaviour Research and Therapy*, *96*, 3–13. <https://doi.org/10.1016/j.brat.2017.04.004>
- Arnaudova, I., Krypotos, A.-M., Effting, M., Kindt, M., & Beckers, T. (2017). Fearing shades of grey: individual differences in fear responding towards generalisation stimuli. *Cognition and Emotion*, *31*, 1181–1196.
<https://doi.org/10.1080/02699931.2016.1204990>

- Asmundson, G. J. G., Stapleton, J. A., & Taylor, S. (2004). Are avoidance and numbing distinct PTSD symptom clusters? *Journal of Traumatic Stress, 17*, 467–475. <https://doi.org/10.1007/s10960-004-5795-7>
- Asmundson, G. J. G., & Taylor, S. (1996). Role of anxiety sensitivity in pain-related fear and avoidance. *Journal of Behavioral Medicine, 19*, 577–586. <https://doi.org/10.1007/BF01904905>
- Aupperle, R. L., Sullivan, S., Melrose, A. J., Paulus, M. P., & Stein, M. B. (2011). A reverse translational approach to quantify approach-avoidance conflict in humans. *Behavioural Brain Research, 225*, 455–463. <https://doi.org/10.1016/j.bbr.2011.08.003>
- Aupperle, R., & Paulus, M. (2010). Neural systems underlying approach and avoidance in anxiety disorders. *Dialogues in Clinical Neuroscience, 12*, 517–531.
- Bach, D. R., Castegnetti, G., Korn, C. W., Gerster, S., Melinscak, F., & Moser, T. (2018). Psychophysiological modeling: current state and future directions. *Psychophysiology, 55*, e13214. <https://doi.org/10.1111/psyp.13209>
- Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment, 13*, 27–45. <https://doi.org/10.1177/1073191105283504>
- Bagiella, E., Sloan, R. P., & Heitjan, D. F. (2000). Mixed-effects models in psychophysiology. *Psychophysiology, 37*, 13–20.
- Bakker, M., van Dijk, A., & Wicherts, J. M. (2012). The rules of the game called psychological science. *Perspectives on Psychological Science, 7*, 543–554. <https://doi.org/10.1177/1745691612459060>
- Balderston, N. L., Vytal, K. E., O'Connell, K., Torrisi, S., Letkiewicz, A., Ernst, M., & Grillon, C. (2017). Anxiety patients show reduced working memory related dlPFC activation during safety and threat. *Depression and Anxiety, 34*, 25–36. <https://doi.org/10.1002/da.22518>
- Ball, T. M., Stein, M. B., & Paulus, M. P. (2014). Toward the application of functional neuroimaging to individualized treatment for anxiety and depression. *Depression and Anxiety, 31*, 920–933. <https://doi.org/10.1002/da.22299>
- Barber, J. P. (2009). Toward a working through of some core conflicts in psychotherapy research. *Psychotherapy Research, 19*, 1–12. <https://doi.org/10.1080/10503300802609680>
- Barlow, D. H., Craske, M. G., Cerny, J. A., & Klosko, J. S. (1989). Behavioral treatment of panic disorder. *Behavior Therapy, 20*, 261–282. [https://doi.org/10.1016/S0005-7894\(89\)80073-5](https://doi.org/10.1016/S0005-7894(89)80073-5)
- Barlow, D. H., Ellard, K. K., Sauer-Zavala, S., Bullis, J. R., & Carl, J. R. (2014). The origins of neuroticism. *Perspectives on Psychological Science, 9*, 481–496.
- Barlow, D. H., Farchione, T. J., Fairholme, C. P., Ellard, K. K., Boisseau, C. L., Allen, L. B., & May, J. T. E. (2010). *Unified Protocol for Transdiagnostic Treatment of Emotional Disorders: Therapist Guide*. Oxford University Press.
- Barlow, D. H., Sauer-Zavala, S., Carl, J. R., Bullis, J. R., & Ellard, K. K. (2013). The nature, diagnosis, and treatment of neuroticism back to the future. *Clinical Psychological Science, 2*, 167702613505532. <https://doi.org/10.1177/2167702613505532>

- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: keep it maximal. *Journal of Memory and Language*, 68. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bartley, C. E., & Roesch, S. C. (2011). Coping with daily stress: the role of conscientiousness. *Personality and Individual Differences*, 50, 79–83. <https://doi.org/10.1016/j.paid.2010.08.027>
- Bastiaens, T., Claes, L., Smits, D., De Clercq, B., De Fruyt, F., Rossi, G., ... De Hert, M. (2016). The construct validity of the Dutch personality inventory for DSM-5 personality disorders (PID-5) in a clinical sample. *Assessment*, 23, 42–51. <https://doi.org/10.1177/1073191115575069>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67. <https://doi.org/10.18637/jss.v067.i01>
- Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: inferential and graphical techniques. *Multivariate Behavioral Research*, 40, 373–400. https://doi.org/10.1207/s15327906mbr4003_5
- Beck, A. T., Davis, D. D., & Freeman, A. (2015). *Cognitive Therapy of Personality Disorders, Third Edition*. Guilford Publications.
- Beck, A. T., Steer, R. A., & Brown, G. K. (1996). *Manual for the Beck Depression Inventory-II*. San Antonio, TX: Psychological Corporation.
- Beckers, T., Krypotos, A.-M., Boddez, Y., Eftting, M., & Kindt, M. (2013). What's wrong with fear conditioning? *Biological Psychology*, 92, 90–96. <https://doi.org/10.1016/j.biopsycho.2011.12.015>
- Behar, E., DiMarco, I. D., Hekler, E. B., Mohlman, J., & Staples, A. M. (2009). Current theoretical models of generalized anxiety disorder (GAD): conceptual review and treatment implications. *Journal of Anxiety Disorders*, 23, 1011–1023. <https://doi.org/10.1016/j.janxdis.2009.07.006>
- Benning, S. D. (2011). Postauricular and superior auricular reflex modulation during emotional pictures and sounds. *Psychophysiology*, 48, 410–414. <https://doi.org/10.1111/j.1469-8986.2010.01071.x>
- Benning, S. D., Patrick, C. J., & Lang, A. R. (2004). Emotional modulation of the post-auricular reflex. *Psychophysiology*, 41, 426–432. <https://doi.org/10.1111/j.1469-8986.00160.x>
- Berkhof, J., & Snijders, T. A. B. (2001). Variance component testing in multilevel models. *Journal of Educational and Behavioral Statistics*, 26, 133–152. <https://doi.org/10.3102/10769986026002133>
- Besnard, A., & Sahay, A. (2016). Adult hippocampal neurogenesis, fear generalization, and stress. *Neuropsychopharmacology*, 41, 24–44. <https://doi.org/10.1038/npp.2015.167>
- Bieling, P. J., Antony, M. M., & Swinson, R. P. (1998). The state--trait anxiety inventory, trait version: structure and content re-examined. *Behaviour Research and Therapy*, 36, 777–788. [https://doi.org/10.1016/S0005-7967\(98\)00023-0](https://doi.org/10.1016/S0005-7967(98)00023-0)
- Bienvenu, O. J., Samuels, J. F., Costa, P. T., Reti, I. M., Eaton, W. W., & Nestadt, G. (2004). Anxiety and depressive disorders and the five-factor model of personality: a higher- and lower-order personality trait investigation in a community sample. *Depression and Anxiety*, 20, 92–97. <https://doi.org/10.1002/da.20026>

- Bilgiç, R., & Sümer, H. C. (2009). Predicting military performance from specific personality measures: a validity study. *International Journal of Selection and Assessment*, 17, 231–238. <https://doi.org/10.1111/j.1468-2389.2009.00465.x>
- Blake, D. D., Weathers, F. W., Nagy, L. M., Kaloupek, D. G., Gusman, F. D., Charney, D. S., & Keane, T. M. (1995). The development of a clinician-administered PTSD scale. *Journal of Traumatic Stress*, 8, 75–90.
- Blechert, J., Michael, T., Vriends, N., Margraf, J., & Wilhelm, F. H. (2007). Fear conditioning in posttraumatic stress disorder: evidence for delayed extinction of autonomic, experiential, and behavioural responses. *Behaviour Research and Therapy*, 45, 2019–2033. <https://doi.org/10.1016/j.brat.2007.02.012>
- Blumenthal, T. D., Cuthbert, B. N., Filion, D. L., Hackley, S., Lipp, O. V., & Van Boxtel, A. (2005). Committee report: guidelines for human startle eyeblink electromyographic studies. *Psychophysiology*, 42, 1–15. <https://doi.org/10.1111/j.1469-8986.2005.00271.x>
- Boddez, Y., Baeyens, F., Luyten, L., Vansteenwegen, D., Hermans, D., & Beckers, T. (2013). Rating data are underrated: validity of us expectancy in human fear conditioning. *Journal of Behavior Therapy and Experimental Psychiatry*, 44, 201–206. <https://doi.org/10.1016/j.jbtep.2012.08.003>
- Boddez, Y., Bennett, M. P., van Esch, S., & Beckers, T. (2016). Bending rules: the shape of the perceptual generalisation gradient is sensitive to inference rules. *Cognition & Emotion*, 1–9. <https://doi.org/10.1080/02699931.2016.1230541>
- Boettcher, H., & Barlow, D. H. (2018). The unique and conditional effects of interoceptive exposure in the treatment of anxiety: a functional analysis. *Behaviour Research and Therapy*. <https://doi.org/10.1016/j.brat.2018.12.002>
- Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J.-S. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. *Trends in Ecology & Evolution*, 24, 127–135. <https://doi.org/10.1016/j.tree.2008.10.008>
- Bonnelle, V., Veromann, K.-R., Burnett Heyes, S., Lo Sterzo, E., Manohar, S., & Husain, M. (2015). Characterization of reward and effort mechanisms in apathy. *Journal of Physiology-Paris*, 109, 16–26. <https://doi.org/10.1016/j.jphysparis.2014.04.002>
- Borkovec, T. D., Alcaine, O. M., & Behar, E. (2004). Avoidance theory of worry and generalized anxiety disorder. In R. G. Heimberg, C. L. Turk, & D. S. Mennin (Eds.), *Generalized anxiety disorder: Advances in research and practice*. (pp. 77–108). New York, NY US: Guilford Press.
- Bouton, M. E., Mineka, S., & Barlow, D. H. (2001). A modern learning theory perspective on the etiology of panic disorder. *Psychological Review*, 108, 4–32. <https://doi.org/10.1037/0033-295X.108.1.4>
- Bouton, M. E., & Swartzentruber, D. (1991). Sources of relapse after extinction in Pavlovian and instrumental learning. *Clinical Psychology Review*, 11, 123–140.
- Boyle, S., Roche, B., Dymond, S., & Hermans, D. (2016). Generalisation of fear and avoidance along a semantic continuum. *Cognition and Emotion*, 30, 340–352. <https://doi.org/10.1080/02699931.2014.1000831>
- Bradford, D. E., Kaye, J. T., & Curtin, J. J. (2014). Not just noise: individual differences in general startle reactivity predict startle response to uncertain and certain threat. *Psychophysiology*, 51, 407–411. <https://doi.org/10.1111/psyp.12193>

- Bradford, D. E., Starr, M. J., Shackman, A. J., & Curtin, J. J. (2015). Empirically based comparisons of the reliability and validity of common quantification approaches for eyeblink startle potentiation in humans. *Psychophysiology*, 52, 1669–1681. <https://doi.org/10.1111/psyp.12545>
- Bradley, M. M., Moulder, B., & Lang, P. J. (2005). When good things go bad: the reflex physiology of defense. *Psychological Science*, 16, 468–473. <https://doi.org/10.1111/j.0956-7976.2005.01558.x>
- Breslau, N. (2009). The epidemiology of trauma, PTSD, and other posttrauma disorders. *Trauma, Violence, & Abuse*, 10, 198–210. <https://doi.org/10.1177/1524838009334448>
- Broman-Fulks, J. J., Urbaniak, A., Bondy, C. L., & Toomey, K. J. (2014). Anxiety sensitivity and risk-taking behavior. *Anxiety, Stress, & Coping*, 27, 619–632. <https://doi.org/10.1080/10615806.2014.896906>
- Bryant, R. A., Creamer, M., O'Donnell, M., Forbes, D., McFarlane, A. C., Silove, D., & Hadzi-Pavlovic, D. (2017). Acute and chronic posttraumatic stress symptoms in the emergence of posttraumatic stress disorder: a network analysis. *JAMA Psychiatry*, 74, 135–142. <https://doi.org/10.1001/jamapsychiatry.2016.3470>
- Büchel, C., & Dolan, R. J. (2000). Classical fear conditioning in functional neuroimaging. *Current Opinion in Neurobiology*, 10, 219–223. [https://doi.org/10.1016/S0959-4388\(00\)00078-7](https://doi.org/10.1016/S0959-4388(00)00078-7)
- Buhr, K., & Dugas, M. J. (2002). The intolerance of uncertainty scale: psychometric properties of the English version. *Behaviour Research and Therapy*, 40, 931–945. [https://doi.org/10.1016/S0005-7967\(01\)00092-4](https://doi.org/10.1016/S0005-7967(01)00092-4)
- Burns, L., & Teesson, M. (2002). Alcohol use disorders comorbid with anxiety, depression and drug use disorders: findings from the Australian national survey of mental health and well being. *Drug and Alcohol Dependence*, 68, 299–307. [https://doi.org/10.1016/S0376-8716\(02\)00220-X](https://doi.org/10.1016/S0376-8716(02)00220-X)
- Buss, K. A., Davis, E. L., Ram, N., & Coccia, M. (2018). Dysregulated fear, social inhibition, and respiratory sinus arrhythmia: a replication and extension. *Child Development*, 89, e214–e228. <https://doi.org/10.1111/cdev.12774>
- Butcher, J. N., & Rouse, S. V. (1996). PERSONALITY: individual differences and clinical assessment. *Annual Review of Psychology*, 47, 87–111. <https://doi.org/10.1146/annurev.psych.47.1.87>
- Bylsma, L. M., Morris, B. H., & Rottenberg, J. (2008). A meta-analysis of emotional reactivity in major depressive disorder. *Clinical Psychology Review*, 28, 676–691. <https://doi.org/10.1016/j.cpr.2007.10.001>
- Byrne, K. A., Silasi-Mansat, C. D., & Worthy, D. A. (2015). Who chokes under pressure? The Big Five personality traits and decision-making under pressure. *Personality and Individual Differences*, 74, 22–28. <https://doi.org/10.1016/j.paid.2014.10.009>
- Bystriksy, A., Craske, M., Maidenberg, E., Vapnik, T., & Shapiro, D. (2000). Autonomic reactivity of panic patients during a CO₂ inhalation procedure. *Depression and Anxiety*, 11, 15–26. [https://doi.org/10.1002/\(SICI\)1520-6394\(2000\)11:1<15::AID-DA3>3.0.CO;2-W](https://doi.org/10.1002/(SICI)1520-6394(2000)11:1<15::AID-DA3>3.0.CO;2-W)
- Cacioppo, J. T., Tassinary, L. G., & Berntson, G. (2007). *Handbook of psychophysiology*. Cambridge University Press.

- Cameron, G., Schlund, M. W., & Dymond, S. (2015). Generalization of socially transmitted and instructed avoidance. *Frontiers in Behavioral Neuroscience*, 9. <https://doi.org/10.3389/fnbeh.2015.00159>
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81.
- Carleton, R. N. (2012). The intolerance of uncertainty construct in the context of anxiety disorders: theoretical and practical perspectives. *Expert Review of Neurotherapeutics*, 12, 937–947. <https://doi.org/10.1586/ern.12.82>
- Carleton, R. N. (2016). Into the unknown: a review and synthesis of contemporary models involving uncertainty. *Journal of Anxiety Disorders*, 39, 30–43. <https://doi.org/10.1016/j.janxdis.2016.02.007>
- Carleton, R. N., Norton, M. A. P. J., & Asmundson, G. J. G. (2007). Fearing the unknown: a short version of the intolerance of uncertainty scale. *Journal of Anxiety Disorders*, 21, 105–117. <https://doi.org/10.1016/j.janxdis.2006.03.014>
- Carter, N. T., Guan, L., Maples, J. L., Williamson, R. L., & Miller, J. D. (2016). The downsides of extreme conscientiousness for psychological well-being: the role of obsessive compulsive tendencies. *Journal of Personality*, 84, 510–522. <https://doi.org/10.1111/jopy.12177>
- Carter, R. M., Hofstötter, C., Tsuchiya, N., & Koch, C. (2003). Working memory and fear conditioning. *Proceedings of the National Academy of Sciences*, 100, 1399–1404. <https://doi.org/10.1073/pnas.0334049100>
- Carver, C. S. (1998). Generalization, adverse events, and development of depressive symptoms. *Journal of Personality*, 66, 607–619. <https://doi.org/10.1111/1467-6494.00026>
- Carver, C. S., & Harmon-Jones, E. (2009). Anger is an approach-related affect: evidence and implications. *Psychological Bulletin*, 135, 183.
- Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: the BIS/BAS scales. *Journal of Personality and Social Psychology*, 67, 319–333. <https://doi.org/10.1037/0022-3514.67.2.319>
- Casa, L. G. D. la, Mena, A., Ruiz-Salas, J. C., Quintero, E., & Papini, M. R. (2018). Reward devaluation disrupts latent inhibition in fear conditioning. *Learning & Behavior*, 46, 49–59. <https://doi.org/10.3758/s13420-017-0282-1>
- Cattell, R. B., & Scheier, I. H. (1961). *The meaning and measurement of neuroticism and anxiety*. Oxford, England: Ronald.
- Çevik, M. Ö. (2014). Habituation, sensitization, and Pavlovian conditioning. *Frontiers in Integrative Neuroscience*, 8, 13. <https://doi.org/10.3389/fnint.2014.00013>
- Cha, J., Greenberg, T., Carlson, J. M., DeDora, D. J., Hajcak, G., & Mujica-Parodi, L. R. (2014). Circuit-wide structural and functional measures predict ventromedial prefrontal cortex fear generalization: implications for generalized anxiety disorder. *The Journal of Neuroscience*, 34, 4043–4053. <https://doi.org/10.1523/JNEUROSCI.3372-13.2014>
- Chase, H. W., Kumar, P., Eickhoff, S. B., & Dombrovski, A. Y. (2015). Reinforcement learning models and their neural correlates: an activation likelihood estimation meta-analysis. *Cognitive, Affective, & Behavioral Neuroscience*, 15, 435–459. <https://doi.org/10.3758/s13415-015-0338-7>

- Chen, J. T.-H., & Lovibond, P. F. (2016). Intolerance of uncertainty is associated with increased threat appraisal and negative affect under ambiguity but not uncertainty. *Behavior Therapy, 47*, 42–53. <https://doi.org/10.1016/j.beth.2015.09.004>
- Chen, S., Yao, N., & Qian, M. (2018). The influence of uncertainty and intolerance of uncertainty on anxiety. *Journal of Behavior Therapy and Experimental Psychiatry, 61*, 60–65. <https://doi.org/10.1016/j.jbtep.2018.06.005>
- Chin, B., Nelson, B. D., Jackson, F., & Hajcak, G. (2016). Intolerance of uncertainty and startle potentiation in relation to different threat reinforcement rates. *International Journal of Psychophysiology, 100*, 1–10. <https://doi.org/10.1016/j.ijpsycho.2015.11.006>
- Choy, Y., Fyer, A. J., & Lipsitz, J. D. (2007). Treatment of specific phobia in adults. *Clinical Psychology Review, 27*, 266–286. <https://doi.org/10.1016/j.cpr.2006.10.002>
- Cisler, J. M., Olatunji, B. O., & Lohr, J. M. (2009). Disgust, fear, and the anxiety disorders: a critical review. *Clinical Psychology Review, 29*, 34–46. <https://doi.org/10.1016/j.cpr.2008.09.007>
- Claes, N., Vlaeyen, J. W. S., & Crombez, G. (2016). Pain in context: cues predicting a reward decrease fear of movement related pain and avoidance behavior. *Behaviour Research and Therapy, 84*, 35–44. <https://doi.org/10.1016/j.brat.2016.07.004>
- Clara, I. P., Cox, B. J., & Enns, M. W. (2003). Hierarchical models of personality and psychopathology: the case of self-criticism, neuroticism, and depression. *Personality and Individual Differences, 35*, 91–99. [https://doi.org/10.1016/S0191-8869\(02\)00143-5](https://doi.org/10.1016/S0191-8869(02)00143-5)
- Clark, D. M., & Wells, A. (1995). A cognitive model of social phobia. *Social Phobia: Diagnosis, Assessment, and Treatment, 41*, 00022–3.
- Clark, L. A., & Watson, D. (1991). Tripartite model of anxiety and depression: psychometric evidence and taxonomic implications. *Journal of Abnormal Psychology, 100*, 316–336. <https://doi.org/10.1037/0021-843X.100.3.316>
- Clark, L. A., Watson, D., & Mineka, S. (1994). Temperament, personality, and the mood and anxiety disorders. *Journal of Abnormal Psychology, 103*, 103–116. <https://doi.org/10.1037/0021-843X.103.1.103>
- Clark, L. A. (2005). Temperament as a unifying basis for personality and psychopathology. *Journal of Abnormal Psychology, 114*, 505–521. <https://doi.org/10.1037/0021-843X.114.4.505>
- Clark, L. A., Cuthbert, B., Lewis-Fernández, R., Narrow, W. E., & Reed, G. M. (2017). Three approaches to understanding and classifying mental disorder: ICD-11, DSM-5, and the national institute of mental health's research domain criteria (RDoC). *Psychological Science in the Public Interest, 18*, 72–145.
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist, 49*, 997–1003. <https://doi.org/10.1037/0003-066X.49.12.997>
- Cohen, S., Kamarck, T., & Mermelstein, R. (1994). Perceived stress scale. *Measuring Stress: A Guide for Health and Social Scientists, 235–283*.
- Contractor, A. A., Armour, C., Shea, M. T., Mota, N., & Pietrzak, R. H. (2016). Latent profiles of DSM-5 PTSD symptoms and the “Big Five” personality traits. *Journal of Anxiety Disorders, 37*, 10–20. <https://doi.org/10.1016/j.janxdis.2015.10.005>

- Contractor, A. A., Roley-Roberts, M. E., Lagdon, S., & Armour, C. (2017). Heterogeneity in patterns of DSM-5 posttraumatic stress disorder and depression symptoms: latent profile analyses. *Journal of Affective Disorders*, 212, 17–24. <https://doi.org/10.1016/j.jad.2017.01.029>
- Conway, C. C., Forbes, M. K., Forbush, K. T., Fried, E. I., Hallquist, M. N., Kotov, R., ... Eaton, N. R. (2019). A hierarchical taxonomy of psychopathology can transform mental health research. *Perspectives on Psychological Science*, 14, 419–436. <https://doi.org/10.1177/1745691618810696>
- Cooper, S. E., Grillon, C., & Lissek, S. (2018). Impaired discriminative fear conditioning during later training trials differentiates generalized anxiety disorder, but not panic disorder, from healthy control participants. *Comprehensive Psychiatry*, 85, 84–93. <https://doi.org/10.1016/j.comppsy.2018.07.001>
- Corr, P. J. (2002). J. A. Gray's reinforcement sensitivity theory: tests of the joint subsystems hypothesis of anxiety and impulsivity. *Personality and Individual Differences*, 33, 511–532. [https://doi.org/10.1016/S0191-8869\(01\)00170-2](https://doi.org/10.1016/S0191-8869(01)00170-2)
- Corr, P. J. (2013). Approach and avoidance behaviour: multiple systems and their interactions. *Emotion Review*, 5, 285–290. <https://doi.org/10.1177/1754073913477507>
- Corr, P. J., & Matthews, G. (2009). The Cambridge handbook of personality psychology, 906.
- Costa, P. T., & McCrae, R. R. (1980). Influence of extraversion and neuroticism on subjective well-being: happy and unhappy people. *Journal of Personality and Social Psychology*, 38, 668.
- Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: the NEO personality inventory. *Psychological Assessment*, 4, 5.
- Costa, P. T., & McCrae, R. R. (1995). Domains and facets: hierarchical personality assessment using the revised NEO personality inventory. *Journal of Personality Assessment*, 64, 21–50. https://doi.org/10.1207/s15327752jpa6401_2
- Craske, M. G., & Barlow, D. H. (1988). A review of the relationship between panic and avoidance. *Clinical Psychology Review*, 8, 667–685. [https://doi.org/10.1016/0272-7358\(88\)90086-4](https://doi.org/10.1016/0272-7358(88)90086-4)
- Craske, M. G., Hermans, D., & Vervliet, B. (2018). State-of-the-art and future directions for extinction as a translational model for fear and anxiety. *Phil. Trans. R. Soc. B*, 373, 20170025. <https://doi.org/10.1098/rstb.2017.0025>
- Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: an inhibitory learning approach. *Behaviour Research and Therapy*, 58, 10–23. <https://doi.org/10.1016/j.brat.2014.04.006>
- Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist*, 12, 671–684. <https://doi.org/10.1037/h0043943>
- Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. *American Psychologist*, 30, 116.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281–302. <https://doi.org/10.1037/h0040957>
- Cuijpers, P., Ebert, D. D., Acarturk, C., Andersson, G., & Cristea, I. A. (2016). Personalized psychotherapy for adult depression: a meta-analytic review. *Behavior Therapy*, 47, 966–980. <https://doi.org/10.1016/j.beth.2016.04.007>

- Cuthbert, B. N., & Insel, T. R. (2013). Toward the future of psychiatric diagnosis: the seven pillars of RDoC. *BMC Medicine*, *11*, 126.
- Davidson, P. O., Payne, R. W., & Sloane, R. B. (1964). Introversion, neuroticism, and conditioning. *The Journal of Abnormal and Social Psychology*, *68*, 136–143. <https://doi.org/10.1037/h0047035>
- Davis, M. (1992). The role of the amygdala in fear-potentiated startle: implications for animal models of anxiety. *Trends in Pharmacological Sciences*, *13*, 35–41. [https://doi.org/10.1016/0165-6147\(92\)90014-W](https://doi.org/10.1016/0165-6147(92)90014-W)
- Dawson, N. V., & Weiss, R. (2012). Dichotomizing continuous variables in statistical analysis: a practice to avoid. *Medical Decision Making*, *32*, 225–226. <https://doi.org/10.1177/0272989X12437605>
- De Cort, K., Griez, E., Büchler, M., & Schruers, K. (2012). The role of “interoceptive” fear conditioning in the development of panic disorder. *Behavior Therapy*, *43*, 203–215. <https://doi.org/10.1016/j.beth.2011.06.005>
- De Cort, K., Schroijen, M., Hurlemann, R., Claassen, S., Hoogenhout, J., Van den Bergh, O., ... Schruers, K. (2017). Modeling the development of panic disorder with interoceptive conditioning. *European Neuropsychopharmacology*, *27*, 59–69. <https://doi.org/10.1016/j.euroneuro.2016.11.001>
- De Fruyt, F., De Clercq, B., De Bolle, M., Wille, B., Markon, K., & Krueger, R. F. (2013). General and maladaptive traits in a five-factor framework for DSM-5 in a university student sample. *Assessment*, *20*, 295–307. <https://doi.org/10.1177/1073191113475808>
- Delgado, M. R., Olsson, A., & Phelps, E. A. (2006). Extending animal models of fear conditioning to humans. *Biological Psychology*, *73*, 39–48. <https://doi.org/10.1016/j.biopsycho.2006.01.006>
- Delgado, M. R. (2009). Avoiding negative outcomes: tracking the mechanisms of avoidance learning in humans during fear conditioning. *Frontiers in Behavioral Neuroscience*, *3*. <https://doi.org/10.3389/neuro.08.033.2009>
- DeLongis, A., & Holtzman, S. (2005). Coping in context: the role of stress, social support, and personality in coping. *Journal of Personality*, *73*, 1633–1656. <https://doi.org/10.1111/j.1467-6494.2005.00361.x>
- Depue, R. A., & Collins, P. F. (1999). Neurobiology of the structure of personality: dopamine, facilitation of incentive motivation, and extraversion. *Behavioral and Brain Sciences*, *22*, 491–517.
- Deschênes, S. S., Dugas, M. J., & Gouin, J.-P. (2016). Intolerance of uncertainty, worry catastrophizing, and heart rate variability during worry-inducing tasks. *Personality and Individual Differences*, *90*, 199–204. <https://doi.org/10.1016/j.paid.2015.11.015>
- DeYoung, C. G. (2006). Higher-order factors of the Big Five in a multi-informant sample. *Journal of Personality and Social Psychology*, *91*, 1138–1151. <https://doi.org/10.1037/0022-3514.91.6.1138>
- DeYoung, C. G. (2015). Cybernetic Big Five theory. *Journal of Research in Personality*, *56*, 33–58. <https://doi.org/10.1016/j.jrp.2014.07.004>
- DeYoung, C. G., Hirsh, J. B., Shane, M. S., Papademetris, X., Rajeevan, N., & Gray, J. R. (2010). Testing predictions from personality neuroscience: brain structure and

- the Big Five. *Psychological Science*, 21, 820–828.
<https://doi.org/10.1177/0956797610370159>
- DeYoung, C. G., & Krueger, R. F. (2018). A cybernetic theory of psychopathology. *Psychological Inquiry*, 29, 117–138.
<https://doi.org/10.1080/1047840X.2018.1513680>
- DeYoung, C. G., Peterson, J. B., & Higgins, D. M. (2005). Sources of openness/intellect: cognitive and neuropsychological correlates of the fifth factor of personality. *Journal of Personality*, 73, 825–858. <https://doi.org/10.1111/j.1467-6494.2005.00330.x>
- DeYoung, C. G., Quilty, L. C., & Peterson, J. B. (2007). Between facets and domains: 10 aspects of the Big Five. *Journal of Personality and Social Psychology*, 93, 880–896. <https://doi.org/10.1037/0022-3514.93.5.880>
- DeYoung, C. G., Shamosh, N. A., Green, A. E., Braver, T. S., & Gray, J. R. (2009). Intellect as distinct from openness: differences revealed by fMRI of working memory. *Journal of Personality and Social Psychology*, 97, 883–892.
<https://doi.org/10.1037/a0016615>
- DeYoung, C. G., Weisberg, Y. J., Quilty, L. C., & Peterson, J. B. (2013). Unifying the aspects of the Big Five, the interpersonal circumplex, and trait affiliation: Big Five and IPC. *Journal of Personality*, 81, 465–475.
<https://doi.org/10.1111/jopy.12020>
- DeYoung, C. G., Carey, B. E., Krueger, R. F., & Ross, S. R. (2016). 10 aspects of the Big Five in the personality inventory for DSM-5. *Personality Disorders*, 7, 113–123.
<https://doi.org/10.1037/per0000170>
- Dibbets, P., Broek, A. van den, & Evers, E. A. T. (2015). Fear conditioning and extinction in anxiety- and depression-prone persons. *Memory*, 23, 350–364.
<https://doi.org/10.1080/09658211.2014.886704>
- Digman, J. M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, 73, 1246–1256. <https://doi.org/10.1037/0022-3514.73.6.1246>
- DiMauro, J., Carter, S., Folk, J. B., & Kashdan, T. B. (2014). A historical review of trauma-related diagnoses to reconsider the heterogeneity of PTSD. *Journal of Anxiety Disorders*, 28, 774–786. <https://doi.org/10.1016/j.janxdis.2014.09.002>
- Dinovo, S. A., & Vasey, M. W. (2011). Reactive and self-regulatory dimensions of temperament: interactive relations with symptoms of general distress and anhedonia. *Journal of Research in Personality*, 45, 430–440.
<https://doi.org/10.1016/j.jrp.2011.05.002>
- Drabant, E. M., Kuo, J. R., Ramel, W., Blechert, J., Edge, M. D., Cooper, J. R., ... Gross, J. J. (2011). Experiential, autonomic, and neural responses during threat anticipation vary as a function of threat intensity and neuroticism. *NeuroImage*, 55, 401–410. <https://doi.org/10.1016/j.neuroimage.2010.11.040>
- Dugas, M. J., Buhr, K., & Ladouceur, R. (2004). The role of intolerance of uncertainty in etiology and maintenance. In R. G. Heimberg, C. L. Turk, & D. S. Mennin (Eds.), *Generalized anxiety disorder: Advances in research and practice* (pp. 143–163). New York, NY, US: Guilford Press.
- Dugas, M. J., & Ladouceur, R. (2000). Treatment of GAD targeting intolerance of uncertainty in two types of worry. *Behavior Modification*, 24, 635–657.
<https://doi.org/10.1177/0145445500245002>

- Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., ... Baas, J. M. P. (2015). Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depression and Anxiety*, 32, 239–253. <https://doi.org/10.1002/da.22353>
- Duits, P., Hofmeijer-Sevink, M. K., Engelhard, I. M., Baas, J. M. P., Ehrismann, W. A. M., & Cath, D. C. (2016). Threat expectancy bias and treatment outcome in patients with panic disorder and agoraphobia. *Journal of Behavior Therapy and Experimental Psychiatry*. <https://doi.org/10.1016/j.jbtep.2016.03.014>
- Duits, P., Richter, J., P, M., Engelhard, I. M., Limberg-Thiesen, A., Heitland, I., ... Cath, D. C. (2017). Enhancing effects of contingency instructions on fear acquisition and extinction in anxiety disorders. *Journal of Abnormal Psychology*, 126, 378–391. <https://doi.org/10.1037/abn0000266>
- Dunsmoor, J. E., Campese, V. D., Ceceli, A. O., LeDoux, J. E., & Phelps, E. A. (2015). Novelty-facilitated extinction: providing a novel outcome in place of an expected threat diminishes recovery of defensive responses. *Biological Psychiatry*, 78, 203–209. <https://doi.org/10.1016/j.biopsych.2014.12.008>
- Dunsmoor, J. E., Kroes, M. C. W., Braren, S. H., & Phelps, E. A. (2017). Threat intensity widens fear generalization gradients. *Behavioral Neuroscience*, 131, 168–175. <https://doi.org/10.1037/bne0000186>
- Dunsmoor, J. E., & Murphy, G. L. (2015). Categories, concepts, and conditioning: how humans generalize fear. *Trends in Cognitive Sciences*. <https://doi.org/10.1016/j.tics.2014.12.003>
- Dunsmoor, J. E., & Paz, R. (2015). Fear generalization and anxiety: behavioral and neural mechanisms. *Biological Psychiatry*, 78, 336–343. <https://doi.org/10.1016/j.biopsych.2015.04.010>
- Dymond, S., Dunsmoor, J. E., Vervliet, B., Roche, B., & Hermans, D. (2014). Fear generalization in humans: systematic review and implications for anxiety disorder research. *Behavior Therapy*. <https://doi.org/10.1016/j.beth.2014.10.001>
- Dymond, S., Schlund, M. W., Roche, B., & Whelan, R. (2014). The spread of fear: symbolic generalization mediates graded threat-avoidance in specific phobia. *The Quarterly Journal of Experimental Psychology*, 67, 247–259. <https://doi.org/10.1080/17470218.2013.800124>
- Eaton, N. R., Krueger, R. F., Markon, K. E., Keyes, K. M., Skodol, A. E., Wall, M., ... Grant, B. F. (2013). The structure and predictive validity of the internalizing disorders. *Journal of Abnormal Psychology*, 122, 86–92. <https://doi.org/10.1037/a0029598>
- Echambadi, R., & Hess, J. D. (2007). Mean-centering does not alleviate collinearity problems in moderated multiple regression models. *Marketing Science*, 26, 438–445.
- Ehlers, A., & Clark, D. M. (2000). A cognitive model of posttraumatic stress disorder. *Behaviour Research and Therapy*, 38, 319–345. [https://doi.org/10.1016/S0005-7967\(99\)00123-0](https://doi.org/10.1016/S0005-7967(99)00123-0)
- Einstein, D. A. (2014). Extension of the transdiagnostic model to focus on intolerance of uncertainty: a review of the literature and implications for treatment. *Clinical Psychology: Science and Practice*, 21, 280–300. <https://doi.org/10.1111/cpsp.12077>

- Elliot, A. J. (2006). The hierarchical model of approach-avoidance motivation. *Motivation and Emotion*, 30, 111–116. <https://doi.org/10.1007/s11031-006-9028-7>
- Elliot, A. J., & Thrash, T. M. (2002). Approach-avoidance motivation in personality: approach and avoidance temperaments and goals. *Journal of Personality and Social Psychology*, 82, 804–818. <https://doi.org/10.1037//0022-3514.82.5.804>
- Enders, C. K., & Tofighi, D. (2007). Centering predictor variables in cross-sectional multilevel models: a new look at an old issue. *Psychological Methods*, 12, 121.
- Endler, N. S. (1977). The role of person-by-situation interactions in personality theory. In I. Č. Užgiris & F. Weizmann (Eds.), *The Structuring of Experience* (pp. 343–369). Springer US. https://doi.org/10.1007/978-1-4615-8786-6_13
- Endler, N. S., Cox, B. J., Parker, J. D., & Michael, R. (1992). Self-reports of depression and state-trait anxiety: evidence for differential assessment. *Journal of Personality and Social Psychology*, 63, 832–838. <https://doi.org/10.1037/0022-3514.63.5.832>
- Endler, N. S., & Kocovski, N. L. (2001). State and trait anxiety revisited. *Journal of Anxiety Disorders*, 15, 231–245. [https://doi.org/10.1016/S0887-6185\(01\)00060-3](https://doi.org/10.1016/S0887-6185(01)00060-3)
- Etkin, M. D., & Wager, Ph. D. (2007). Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *American Journal of Psychiatry*, 164, 1476–1488. <https://doi.org/10.1176/appi.ajp.2007.07030504>
- Eysenck, H. J. (1962). Conditioning and personality. *British Journal of Psychology*, 53, 299–305. <https://doi.org/10.1111/j.2044-8295.1962.tb00835.x>
- Eysenck, H. J. (1979). The conditioning model of neurosis. *Behavioral and Brain Sciences*, 2, 155–166. <https://doi.org/10.1017/S0140525X00061653>
- Eysenck, H. J. (1950). *Dimensions of Personality*. Transaction Publishers.
- Eysenck, H. J., & Eysenck, M. W. (1987). *Personality and individual differences*. Plenum New York, NY.
- Eysenck, H. J. (1970). *The biological basis of personality*. Thomas.
- Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: attentional control theory. *Emotion*, 7, 336–353. <https://doi.org/10.1037/1528-3542.7.2.336>
- Feingold, A. (1994). Gender differences in personality: a meta-analysis. *Psychological Bulletin*, 116, 429.
- FeldmanHall, O., Dunsmoor, J. E., Tompary, A., Hunter, L. E., Todorov, A., & Phelps, E. A. (2018). Stimulus generalization as a mechanism for learning to trust. *Proceedings of the National Academy of Sciences*, 201715227. <https://doi.org/10.1073/pnas.1715227115>
- Fergus, T. A., & Bardeen, J. R. (2013). Anxiety sensitivity and intolerance of uncertainty: evidence of incremental specificity in relation to health anxiety. *Personality and Individual Differences*, 55, 640–644. <https://doi.org/10.1016/j.paid.2013.05.016>
- Fernando, A. B. P., Urcelay, G. P., Mar, A. C., Dickinson, A., & Robbins, T. W. (2014). Safety signals as instrumental reinforcers during free-operant avoidance. *Learning & Memory*, 21, 488–497. <https://doi.org/10.1101/lm.034603.114>
- Few, L. R., Miller, J. D., Rothbaum, A. O., Meller, S., Maples, J., Terry, D. P., ... MacKillop, J. (2013). Examination of the section III DSM-5 diagnostic system for

- personality disorders in an outpatient clinical sample. *Journal of Abnormal Psychology*, 122, 1057–1069. <https://doi.org/10.1037/a0034878>
- Fiedler, K. (2011). Voodoo correlations are everywhere—not only in neuroscience. *Perspectives on Psychological Science*, 6, 163–171. <https://doi.org/10.1177/1745691611400237>
- Field, A. P. (2006). Is conditioning a useful framework for understanding the development and treatment of phobias? *Clinical Psychology Review*, 26, 857–875. <https://doi.org/10.1016/j.cpr.2005.05.010>
- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (2002). *Structured Clinical Interview for DSM-IV-TR Axis I Disorders, Research Version, Patient Edition (SCID-I/P)*. New York: New York State Psychiatric Institute, Biometrics Research.
- Fischer, R., & Milfont, T. L. (2010). Standardization in psychological research. *International Journal of Psychological Research*, 3.
- Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from different sources. *The Journal of Abnormal and Social Psychology*, 44, 329–344. <https://doi.org/10.1037/h0057198>
- Flores, A., López, F. J., Vervliet, B., & Cobos, P. L. (2018). Intolerance of uncertainty as a vulnerability factor for excessive and inflexible avoidance behavior. *Behaviour Research and Therapy*, 104, 34–43. <https://doi.org/10.1016/j.brat.2018.02.008>
- Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: exposure to corrective information. *Psychological Bulletin*, 99, 20–35. <https://doi.org/10.1037/0033-2909.99.1.20>
- Foa, E. B., & McLean, C. P. (2016). The efficacy of exposure therapy for anxiety-related disorders and its underlying mechanisms: the case of OCD and PTSD. *Annual Review of Clinical Psychology*, 12, 1–28. <https://doi.org/10.1146/annurev-clinPSY-021815-093533>
- Foa, E. B., Steketee, G., & Rothbaum, B. O. (1989). Behavioral/cognitive conceptualizations of post-traumatic stress disorder. *Behavior Therapy*, 20, 155–176. [https://doi.org/10.1016/S0005-7894\(89\)80067-X](https://doi.org/10.1016/S0005-7894(89)80067-X)
- Foa, E., Hembree, E., & Rothbaum, B. O. (2007). *Prolonged Exposure Therapy for PTSD: Emotional Processing of Traumatic Experiences Therapist Guide*. Oxford University Press, USA.
- Forbes, M. K., Kotov, R., Ruggero, C. J., Watson, D., Zimmerman, M., & Krueger, R. F. (2017). Delineating the joint hierarchical structure of clinical and personality disorders in an outpatient psychiatric sample. *Comprehensive Psychiatry*, 79, 19–30. <https://doi.org/10.1016/j.comppsy.2017.04.006>
- Forsyth, J. P., Palav, A., & Duff, K. (1999). The absence of relation between anxiety sensitivity and fear conditioning using 20% versus 13% CO₂-enriched air as unconditioned stimuli. *Behaviour Research and Therapy*, 37, 143–153. [https://doi.org/10.1016/S0005-7967\(98\)00113-2](https://doi.org/10.1016/S0005-7967(98)00113-2)
- Fossati, A., Krueger, R. F., Markon, K. E., Borroni, S., & Maffei, C. (2013). Reliability and validity of the personality inventory for DSM-5 (PID-5): predicting DSM-IV personality disorders and psychopathy in community-dwelling Italian adults. *Assessment*, 20, 689–708. <https://doi.org/10.1177/1073191113504984>

- Fox, A. S., & Shackman, A. J. (2017). The central extended amygdala in fear and anxiety: closing the gap between mechanistic and neuroimaging research. *Neuroscience Letters*, 693, 58–67. <https://doi.org/10.1016/j.neulet.2017.11.056>
- Franks, C. M. (1956). Conditioning and personality: a study of normal and neurotic subjects. *The Journal of Abnormal and Social Psychology*, 52, 143–150. <https://doi.org/10.1037/h0047250>
- Friedrich, R. J. (1982). In defense of multiplicative terms in multiple regression equations. *American Journal of Political Science*, 26, 797–833. <https://doi.org/10.2307/2110973>
- Fullana, M. A., Vilagut, G., Rojas-Farreras, S., Mataix-Cols, D., de Graaf, R., Demyttenaere, K., ... Alonso, J. (2010). Obsessive–compulsive symptom dimensions in the general population: results from an epidemiological study in six European countries. *Journal of Affective Disorders*, 124, 291–299. <https://doi.org/10.1016/j.jad.2009.11.020>
- Funayama, T., Furukawa, T. A., Nakano, Y., Noda, Y., Ogawa, S., Watanabe, N., ... Noguchi, Y. (2013). In-situation safety behaviors among patients with panic disorder: descriptive and correlational study. *Psychiatry and Clinical Neurosciences*, 67, 332–339. <https://doi.org/10.1111/pcn.12061>
- Garbusow, M., Schad, D. J., Sebold, M., Friedel, E., Bernhardt, N., Koch, S. P., ... Heinz, A. (2016). Pavlovian-to-instrumental transfer effects in the nucleus accumbens relate to relapse in alcohol dependence. *Addiction Biology*, 21, 719–731. <https://doi.org/10.1111/adb.12243>
- Garofalo, S., & Robbins, T. W. (2017). Triggering avoidance: dissociable influences of aversive Pavlovian conditioned stimuli on human instrumental behavior. *Frontiers in Behavioral Neuroscience*, 11. <https://doi.org/10.3389/fnbeh.2017.00063>
- Gaylin, W. (1979). *Feelings: Our vital signs*. Harper & Row New York.
- Gazendam, F. J., Kamphuis, J. H., Eigenhuis, A., Huizenga, H. M. H., Soeter, M., Bos, M. G. N., ... Kindt, M. (2014). Personality predicts individual variation in fear learning: a multilevel growth modeling approach. *Clinical Psychological Science*. <https://doi.org/10.1177/2167702614535914>
- Gazendam, F. J., Kamphuis, J. H., & Kindt, M. (2013). Deficient safety learning characterizes high trait anxious individuals. *Biological Psychology*, 92, 342–352. <https://doi.org/10.1016/j.biopsycho.2012.11.006>
- Geller, D. A., McGuire, J. F., Orr, S. P., Pine, D. S., Britton, J. C., Small, B. J., ... Storch, E. A. (2017). Fear conditioning and extinction in pediatric obsessive-compulsive disorder. *Annals of Clinical Psychiatry*, 29, 17–26.
- Gelman, A., & Hill, J. (2006). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.
- Gelman, A., Hill, J., & Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *Journal of Research on Educational Effectiveness*, 5, 189–211. <https://doi.org/10.1080/19345747.2011.618213>
- Gershman, S. J., & Daw, N. D. (2017). Reinforcement learning and episodic memory in humans and animals: an integrative framework. *Annual Review of Psychology*, 68, 101–128. <https://doi.org/10.1146/annurev-psych-122414-033625>

- Gillan, C. M., Apergis-Schoute, A. M., Morein-Zamir, S., Urcelay, G. P., Sule, A., Fineberg, N. A., ... Robbins, T. W. (2014). Functional neuroimaging of avoidance habits in obsessive-compulsive disorder. *American Journal of Psychiatry*, 172, 284–293. <https://doi.org/10.1176/appi.ajp.2014.14040525>
- Gillan, C. M., Morein-Zamir, S., Urcelay, G. P., Sule, A., Voon, V., Apergis-Schoute, A. M., ... Robbins, T. W. (2014). Enhanced avoidance habits in obsessive-compulsive disorder. *Biological Psychiatry*, 75, 631–638. <https://doi.org/10.1016/j.biopsych.2013.02.002>
- Ginat-Frolich, R., Klein, Z., Katz, O., & Shechner, T. (2017). A novel perceptual discrimination training task: reducing fear overgeneralization in the context of fear learning. *Behaviour Research and Therapy*, 93, 29–37. <https://doi.org/10.1016/j.brat.2017.03.010>
- Glimcher, P. W., & Rustichini, A. (2004). Neuroeconomics: the consilience of brain and decision. *Science*, 306, 447–452. <https://doi.org/10.1126/science.1102566>
- Goldberg, L. R. (1981). Language and individual differences: the search for universals in personality lexicons. *Review of Personality and Social Psychology*, 2, 141–165.
- Goldberg, L. R. (1990). An alternative “description of personality”: the big-five factor structure. *Journal of Personality and Social Psychology*, 14.
- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. *Personality Psychology in Europe*, 7, 7–28.
- Goldstein, H., Browne, W., & Rasbash, J. (2002). Partitioning variation in multilevel models. *Understanding Statistics*, 1, 223–231. https://doi.org/10.1207/S15328031US0104_02
- Gomez, R., Cooper, A., & Gomez, A. (2005). An item response theory analysis of the carver and white (1994) BIS/BAS scales. *Personality and Individual Differences*, 39, 1093–1103. <https://doi.org/10.1016/j.paid.2005.03.015>
- Gore, W. L., & Widiger, T. A. (2013). The DSM-5 dimensional trait model and five-factor models of general personality. *Journal of Abnormal Psychology*, 122, 816.
- Grant, B. F., Hasin, D. S., Stinson, F. S., Dawson, D. A., Goldstein, R. B., Smith, S., ... Saha, T. D. (2006). The epidemiology of DSM-IV panic disorder and agoraphobia in the united states: results from the national epidemiologic survey on alcohol and related conditions. *The Journal of Clinical Psychiatry*, 67, 363–374. <https://doi.org/10.4088/JCP.v67n0305>
- Gray, J. A. (1972). The psychophysiological nature of introversion-extraversion: a modification of Eysenck’s theory. *Biological Bases of Individual Behavior*, 182–205.
- Gray, J. A. (1990). Brain systems that mediate both emotion and cognition. *Cognition and Emotion*, 4, 269–288. <https://doi.org/10.1080/02699939008410799>
- Greenberg, T., Carlson, J. M., Cha, J., Hajcak, G., & Mujica-Parodi, L. R. (2013). Ventromedial prefrontal cortex reactivity is altered in generalized anxiety disorder during fear generalization. *Depression and Anxiety*, 30, 242–250. <https://doi.org/10.1002/da.22016>
- Greenwald, M. K., Bradley, M. M., Cuthbert, B. N., & Lang, P. J. (1998). Startle potentiation: shock sensitization, aversive learning, and affective picture

- modulation. *Behavioral Neuroscience*, 112, 1069–1079.
<https://doi.org/10.1037/0735-7044.112.5.1069>
- Grillon, C., & Baas, J. (2002). Comments on the use of the startle reflex in psychopharmacological challenges: impact of baseline startle on measurement of fear-potentiated startle. *Psychopharmacology*, 164, 236–238.
- Grillon, C., Ameli, R., Woods, S. W., Merikangas, K., & Davis, M. (1991). Fear-potentiated startle in humans: effects of anticipatory anxiety on the acoustic blink reflex. *Psychophysiology*, 28, 588–595.
- Grillon, C., & Baas, J. (2003). A review of the modulation of the startle reflex by affective states and its application in psychiatry. *Clinical Neurophysiology*, 114, 1557–1579. [https://doi.org/10.1016/S1388-2457\(03\)00202-5](https://doi.org/10.1016/S1388-2457(03)00202-5)
- Groman, S. M., James, A. S., Seu, E., Tran, S., Clark, T. A., Harpster, S. N., ... Jentsch, J. D. (2014). In the blink of an eye: relating positive-feedback sensitivity to striatal dopamine D2-like receptors through blink rate. *Journal of Neuroscience*, 34, 14443–14454. <https://doi.org/10.1523/JNEUROSCI.3037-14.2014>
- Grös, D. F., Antony, M. M., Simms, L. J., & McCabe, R. E. (2007). Psychometric properties of the state-trait inventory for cognitive and somatic anxiety (STICSA): comparison to the state-trait anxiety inventory (STAI). *Psychological Assessment*, 19, 369–381. <https://doi.org/10.1037/1040-3590.19.4.369>
- Gross, J. J. (2011). *Handbook of Emotion Regulation, First Edition*. Guilford Press.
- Grupe, D. W., & Nitschke, J. B. (2013). Uncertainty and anticipation in anxiety: an integrated neurobiological and psychological perspective. *Nature Reviews Neuroscience*, 14, 488–501. <https://doi.org/10.1038/nrn3524>
- Grupe, D. W., & Nitschke, J. B. (2011). Uncertainty is associated with biased expectancies and heightened responses to aversion. *Emotion (Washington, D.C.)*, 11, 413–424. <https://doi.org/10.1037/a0022583>
- Grupe, D. W. (2017). Decision-making in anxiety and its disorders. In *Decision Neuroscience* (pp. 327–338). Elsevier. <https://doi.org/10.1016/B978-0-12-805308-9.00026-9>
- Gueorguieva, R., & Krystal, J. H. (2004). Move over ANOVA: progress in analyzing repeated-measures data and its reflection in papers published in the archives of general psychiatry. *Archives of General Psychiatry*, 61, 310–317.
- Gutner, C. A., Gallagher, M. W., Baker, A. S., Sloan, D. M., & Resick, P. A. (2015). Time course of treatment dropout in cognitive-behavioral therapies for posttraumatic stress disorder. *Psychological Trauma: Theory, Research, Practice, and Policy*, No Pagination Specified. <https://doi.org/10.1037/tra0000062>
- Haaker, J., Lonsdorf, T. B., Schümann, D., Menz, M., Brassen, S., Bunzeck, N., ... Kalisch, R. (2015). Deficient inhibitory processing in trait anxiety: evidence from context-dependent fear learning, extinction recall and renewal. *Biological Psychology*, 111, 65–72. <https://doi.org/10.1016/j.biopsycho.2015.07.010>
- Haddad, A. D. M., Pritchett, D., Lissek, S., & Lau, J. Y. F. (2012). Trait anxiety and fear responses to safety cues: stimulus generalization or sensitization? *Journal of Psychopathology and Behavioral Assessment*, 34, 323–331. <https://doi.org/10.1007/s10862-012-9284-7>
- Harkness, A. R., & Lilienfeld, S. O. (1997). Individual differences science for treatment planning: personality traits. *Psychological Assessment*, 9, 349.

- Harmon-Jones, E., Peterson, C., Gable, P. A., & Harmon-Jones, C. (2008). Anger and approach-avoidance motivation. In *Handbook of approach and avoidance motivation* (pp. 399–413). New York, NY, US: Psychology Press.
- Harrison, X. A., Donaldson, L., Correa-Cano, M. E., Evans, J., Fisher, D. N., Goodwin, C. E. D., ... Inger, R. (2018). A brief introduction to mixed effects modelling and multi-model inference in ecology. *PeerJ*, 6. <https://doi.org/10.7717/peerj.4794>
- Haslam, N., Williams, B. J., Kyrios, M., McKay, D., & Taylor, S. (2005). Subtyping obsessive-compulsive disorder: a taxometric analysis. *Behavior Therapy*, 36, 381–391. [https://doi.org/10.1016/S0005-7894\(05\)80120-0](https://doi.org/10.1016/S0005-7894(05)80120-0)
- Hassoulas, A., McHugh, L., & Reed, P. (2014). Avoidance and behavioural flexibility in obsessive compulsive disorder. *Journal of Anxiety Disorders*, 28, 148–153. <https://doi.org/10.1016/j.janxdis.2013.05.002>
- Hauck Jr, W. W., & Donner, A. (1977). Wald's test as applied to hypotheses in logit analysis. *Journal of the American Statistical Association*, 72, 851–853.
- Hayes, A. F. (2018). Partial, conditional, and moderated moderated mediation: quantification, inference, and interpretation. *Communication Monographs*, 85, 4–40. <https://doi.org/10.1080/03637751.2017.1352100>
- Hayes, A. F., Glynn, C. J., & Hude, M. E. (2012). Cautions regarding the interpretation of regression coefficients and hypothesis tests in linear models with interactions. *Communication Methods and Measures*, 6, 1–11. <https://doi.org/10.1080/19312458.2012.651415>
- Hayes, S. C. (2002). Acceptance, mindfulness, and science. *Clinical Psychology: Science and Practice*, 9, 101–106. <https://doi.org/10.1093/clipsy.9.1.101>
- Hayes, S. C. (2004). Acceptance and commitment therapy, relational frame theory, and the third wave of behavioral and cognitive therapies. *Behavior Therapy*, 35, 639–665. [https://doi.org/10.1016/S0005-7894\(04\)80013-3](https://doi.org/10.1016/S0005-7894(04)80013-3)
- Hayes, S. C., Strosahl, K. D., & Wilson, K. G. (2011). *Acceptance and Commitment Therapy, Second Edition: The Process and Practice of Mindful Change*. Guilford Press.
- Hayes, S. C., Wilson, K. G., Gifford, E. V., Follette, V. M., & Strosahl, K. (1996). Experiential avoidance and behavioral disorders: a functional dimensional approach to diagnosis and treatment. *Journal of Consulting and Clinical Psychology*, 64, 1152–1168. <https://doi.org/10.1037/0022-006X.64.6.1152>
- Helwig, N. E. (2017). Regression with ordered predictors via ordinal smoothing splines. *Frontiers in Applied Mathematics and Statistics*, 3. <https://doi.org/10.3389/fams.2017.00015>
- Hermann, C., Ziegler, S., Birbaumer, N., & Flor, H. (2002). Psychophysiological and subjective indicators of aversive Pavlovian conditioning in generalized social phobia. *Biological Psychiatry*, 52, 328–337. [https://doi.org/10.1016/S0006-3223\(02\)01385-9](https://doi.org/10.1016/S0006-3223(02)01385-9)
- Hermans, D., Craske, M. G., Mineka, S., & Lovibond, P. F. (2006). Extinction in human fear conditioning. *Biological Psychiatry*, 60, 361–368. <https://doi.org/10.1016/j.biopsych.2005.10.006>
- Hewitt, P. L., & Flett, G. L. (2007). When does conscientiousness become perfectionism? Traits, self-presentation styles, and cognitions suggest a persistent psychopathology. *Current Psychiatry*, 6, 49.

- Hickman, K. L., Stromme, C., & Lippman, L. G. (1998). Learned industriousness: replication in principle. *The Journal of General Psychology*, 125, 213–217.
- Hodgson, R., & Rachman, S. (1974). II. Desynchrony in measures of fear. *Behaviour Research and Therapy*, 12, 319–326. [https://doi.org/10.1016/0005-7967\(74\)90006-0](https://doi.org/10.1016/0005-7967(74)90006-0)
- Hoffman, L., & Rovine, M. J. (2007). Multilevel models for the experimental psychologist: foundations and illustrative examples. *Behavior Research Methods*, 39, 101–117. <https://doi.org/10.3758/BF03192848>
- Hofmann, S. G. (2007). Cognitive factors that maintain social anxiety disorder: a comprehensive model and its treatment implications. *Cognitive Behaviour Therapy*, 36, 193–209. <https://doi.org/10.1080/16506070701421313>
- Hofmann, S. G. (2008). Cognitive processes during fear acquisition and extinction in animals and humans: implications for exposure therapy of anxiety disorders. *Clinical Psychology Review*, 28, 199–210. <https://doi.org/10.1016/j.cpr.2007.04.009>
- Hofstee, W. K., de Raad, B., & Goldberg, L. R. (1992). Integration of the Big Five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology*, 63, 146–163. <https://doi.org/10.1037/0022-3514.63.1.146>
- Holroyd, C. B., Hajcak, G., & Larsen, J. T. (2006). The good, the bad and the neutral: electrophysiological responses to feedback stimuli. *Brain Research*, 1105, 93–101. <https://doi.org/10.1016/j.brainres.2005.12.015>
- Hopwood, C. J., Kotov, R., Krueger, R. F., Watson, D., Widiger, T. A., Althoff, R. R., ... Zimmermann, J. (2018). The time has come for dimensional personality disorder diagnosis. *Personality and Mental Health*, 12, 82–86. <https://doi.org/10.1002/pmh.1408>
- Hopwood, C. J., Schade, N., Krueger, R. F., Wright, A. G. C., & Markon, K. E. (2013). Connecting DSM-5 personality traits and pathological beliefs: toward a unifying model. *Journal of Psychopathology and Behavioral Assessment*, 35, 162–172. <https://doi.org/10.1007/s10862-012-9332-3>
- Hopwood, C. J., Thomas, K. M., Markon, K. E., Wright, A. G. C., & Krueger, R. F. (2012). DSM-5 personality traits and DSM-IV personality disorders. *Journal of Abnormal Psychology*, 121, 424–432. <https://doi.org/10.1037/a0026656>
- Hopwood, C. J., Wright, A. G., Krueger, R. F., Schade, N., Markon, K. E., & Morey, L. C. (2013). DSM-5 pathological personality traits and the personality assessment inventory. *Assessment*, 20, 269–285.
- Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398). John Wiley & Sons.
- Hox, J. J. (2013). Multilevel regression and multilevel structural equation modeling. *The Oxford Handbook of Quantitative Methods*, 2, 281–294.
- Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications*. Routledge.
- Hsu, K. J., Beard, C., Rifkin, L., Dillon, D. G., Pizzagalli, D. A., & Björgvinsson, T. (2015). Transdiagnostic mechanisms in depression and anxiety: the role of rumination and attentional control. *Journal of Affective Disorders*, 188, 22–27. <https://doi.org/10.1016/j.jad.2015.08.008>

- Hull, C. L. (1935). The conflicting psychologies of learning—a way out. *Psychological Review*, 42, 491.
- Hundt, N. E., Barrera, T. L., Arney, J., & Stanley, M. A. (2017). “It’s worth it in the end”: veterans’ experiences in prolonged exposure and cognitive processing therapy. *Cognitive and Behavioral Practice*, 24, 50–57. <https://doi.org/10.1016/j.cbpra.2016.02.003>
- Hunt, C., Cooper, S. E., Hartnell, M. P., & Lissek, S. (2017). Distraction/suppression and distress endurance diminish the extent to which generalized conditioned fear is associated with maladaptive behavioral avoidance. *Behaviour Research and Therapy*, 96, 90–105. <https://doi.org/10.1016/j.brat.2017.04.013>
- Hunt, C., Cooper, S. E., Hartnell, M. P., & Lissek, S. (2019). Anxiety sensitivity and intolerance of uncertainty facilitate associations between generalized Pavlovian fear and maladaptive avoidance decisions. *Journal of Abnormal Psychology*. <https://doi.org/10.1037/abn0000422>
- Hunt, C., Degeneffe, N., Bixby, J., Fleig, R., & Lissek, S. (2018, April). *OCD is associated with increased fear-motivated avoidance of low probability, high intensity threats: evidence from a college sample*. Presented at the Anxiety and Depression Association of America, Washington, DC.
- Hur, J., Jordan, A. D., Berenbaum, H., & Dolcos, F. (2016). Emotion–attention interactions in fear conditioning: moderation by executive load, neuroticism, and awareness. *Biological Psychology*, 121, 213–220.
- Iacobucci, D., Posavac, S. S., Kardes, F. R., Schneider, M., & Popovich, D. (2015). *Toward a More Nuanced Understanding of the Statistical Properties of a Median Split* (SSRN Scholarly Paper No. ID 2663427). Rochester, NY: Social Science Research Network. Retrieved from <https://papers.ssrn.com/abstract=2663427>
- Indovina, I., Robbins, T. W., Núñez-Elizalde, A. O., Dunn, B. D., & Bishop, S. J. (2011). Fear-conditioning mechanisms associated with trait vulnerability to anxiety in humans. *Neuron*, 69, 563–571. <https://doi.org/10.1016/j.neuron.2010.12.034>
- Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., ... Wang, P. (2010). Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *American Journal of Psychiatry*, 167, 748–751. <http://dx.doi.org/10.1176/appi.ajp.2010.09091379>
- Irwin, J. R., & McClelland, G. H. (2003). Negative consequences of dichotomizing continuous predictor variables. *Journal of Marketing Research*, 40, 366–371.
- Jackson, J. J., Wood, D., Bogg, T., Walton, K. E., Harms, P. D., & Roberts, B. W. (2010). What do conscientious people do? Development and validation of the behavioral indicators of conscientiousness (BIC). *Journal of Research in Personality*, 44, 501–511. <https://doi.org/10.1016/j.jrp.2010.06.005>
- Jang, K. L., Livesley, W. J., Angleitner, A., Riemann, R., & Vernon, P. A. (2002). Genetic and environmental influences on the covariance of facets defining the domains of the five-factor model of personality. *Personality and Individual Differences*, 33, 83–101. [https://doi.org/10.1016/S0191-8869\(01\)00137-4](https://doi.org/10.1016/S0191-8869(01)00137-4)
- Jebb, A. T., Parrigon, S., & Woo, S. E. (2017). Exploratory data analysis as a foundation of inductive research. *Human Resource Management Review*, 27, 265–276. <https://doi.org/10.1016/j.hrmr.2016.08.003>

- Jensen-Campbell, L. A., Rosselli, M., Workman, K. A., Santisi, M., Rios, J. D., & Bojan, D. (2002). Agreeableness, conscientiousness, and effortful control processes. *Journal of Research in Personality*, 36, 476–489. [https://doi.org/10.1016/S0092-6566\(02\)00004-1](https://doi.org/10.1016/S0092-6566(02)00004-1)
- Jiang, D., Zhang, D., Chen, Y., He, Z., Gao, Q., Gu, R., & Xu, P. (2018). Trait anxiety and probabilistic learning: behavioral and electrophysiological findings. *Biological Psychology*, 132, 17–26. <https://doi.org/10.1016/j.biopsycho.2017.10.010>
- John, O. P., & Srivastava, S. (1999). The Big Five trait taxonomy: history, measurement, and theoretical perspectives. *Handbook of Personality: Theory and Research*, 2, 102–138.
- Johns, L. C., & van Os, J. (2001). The continuity of psychotic experiences in the general population. *Clinical Psychology Review*, 21, 1125–1141. [https://doi.org/10.1016/S0272-7358\(01\)00103-9](https://doi.org/10.1016/S0272-7358(01)00103-9)
- Johnson, P. C. D. (2014). Extension of Nakagawa & Schielzeth's R2GLMM to random slopes models. *Methods in Ecology and Evolution*, 5, 944–946. <https://doi.org/10.1111/2041-210X.12225>
- Johnson, S. L., Turner, R. J., & Iwata, N. (2003). BIS/BAS levels and psychiatric disorder: an epidemiological study. *Journal of Psychopathology and Behavioral Assessment*, 25, 25–36. <https://doi.org/10.1023/A:1022247919288>
- Joiner Jr, T. E., Steer, R. A., Beck, A. T., Schmidt, N. B., Rudd, M. D., & Catanzaro, S. J. (1999). Physiological hyperarousal: construct validity of a central aspect of the tripartite model of depression and anxiety. *Journal of Abnormal Psychology*, 108, 290.
- Jones, M. C. (1924). A laboratory study of fear: the case of Peter. *The Journal of Genetic Psychology*, 31, 308–315.
- Jorm, A. F., Christensen, H., Henderson, A. S., Jacomb, P. A., Korten, A. E., & Rodgers, B. (2000). Predicting anxiety and depression from personality: is there a synergistic effect of neuroticism and extraversion? *Journal of Abnormal Psychology*, 5.
- Jovanovic, T., Norrholm, S. D., Blanding, N. Q., Davis, M., Duncan, E., Bradley, B., & Ressler, K. J. (2010). Impaired fear inhibition is a biomarker of PTSD but not depression. *Depression and Anxiety*, 27, 244–251. <https://doi.org/10.1002/da.20663>
- Jylhä, P., & Isometsä, E. (2006). The relationship of neuroticism and extraversion to symptoms of anxiety and depression in the general population. *Depression and Anxiety*, 23, 281–289. <https://doi.org/10.1002/da.20167>
- Kaczurkin, A. N., Burton, P. C., Chazin, S. M., Manbeck, A. B., Espensen-Sturges, T., Cooper, S. E., ... Lissek, S. (2016). Neural substrates of overgeneralized conditioned fear in PTSD. *American Journal of Psychiatry*, 174, 125–134. <https://doi.org/10.1176/appi.ajp.2016.15121549>
- Kaczurkin, A. N., & Lissek, S. (2013). Generalization of conditioned fear and obsessive-compulsive traits. *Journal of Psychology & Psychotherapy*, 7, 3.
- Kalish, H. I. (1969). Stimulus generalization. *Learning: Processes*, 20.
- Kashdan, T. B., Barrios, V., Forsyth, J. P., & Steger, M. F. (2006). Experiential avoidance as a generalized psychological vulnerability: comparisons with coping

- and emotion regulation strategies. *Behaviour Research and Therapy*, 44, 1301–1320. <https://doi.org/10.1016/j.brat.2005.10.003>
- Kashdan, T. B., & Rottenberg, J. (2010). Psychological flexibility as a fundamental aspect of health. *Clinical Psychology Review*, 30, 865–878. <https://doi.org/10.1016/j.cpr.2010.03.001>
- Kaufman, J., & Charney, D. (2000). Comorbidity of mood and anxiety disorders. *Depression and Anxiety*, 12, 69–76. [https://doi.org/10.1002/1520-6394\(2000\)12:1+<69::AID-DA9>3.0.CO;2-K](https://doi.org/10.1002/1520-6394(2000)12:1+<69::AID-DA9>3.0.CO;2-K)
- Kaye, J. T., Bradford, D. E., & Curtin, J. J. (2016). Psychometric properties of startle and corrugator response in NPU, affective picture viewing, and resting state tasks. *Psychophysiology*, 53, 1241–1255. <https://doi.org/10.1111/psyp.12663>
- Kelly, D., & Martin, I. (1969). Autonomic reactivity, eyelid conditioning and their relationship to neuroticism and extraversion. *Behaviour Research and Therapy*, 7, 233–244. [https://doi.org/10.1016/0005-7967\(69\)90002-3](https://doi.org/10.1016/0005-7967(69)90002-3)
- Kennedy, B. L., Schwab, J. J., Morris, R. L., & Beldia, G. (2001). Assessment of state and trait anxiety in subjects with anxiety and depressive disorders. *Psychiatric Quarterly*, 72, 263–276. <https://doi.org/10.1023/A:1010305200087>
- Kessels, L. T. E., Ruiter, R. A. C., Wouters, L., & Jansma, B. M. (2014). Neuroscientific evidence for defensive avoidance of fear appeals. *International Journal of Psychology*, 49, 80–88. <https://doi.org/10.1002/ijop.12036>
- Kessler, R. C., Chiu, W. T., Jin, R., Ruscio, A. M., Shear, K., & Walters, E. E. (2006). The epidemiology of panic attacks, panic disorder, and agoraphobia in the national comorbidity survey replication. *Archives of General Psychiatry*, 63, 415–424. <https://doi.org/10.1001/archpsyc.63.4.415>
- Kessler, R. C., Ruscio, A. M., Shear, K., & Wittchen, H.-U. (2009). Epidemiology of anxiety disorders. In M. B. Stein & T. Steckler (Eds.), *Behavioral Neurobiology of Anxiety and Its Treatment* (pp. 21–35). Springer Berlin Heidelberg. Retrieved from http://link.springer.com.ezp2.lib.umn.edu/chapter/10.1007/7854_2009_9
- Khawaja, N. G., & Yu, L. N. H. (2010). A comparison of the 27-item and 12-item intolerance of uncertainty scales. *Clinical Psychologist*, 14, 97–106. <https://doi.org/10.1080/13284207.2010.502542>
- Kim, H., Shimojo, S., & O'Doherty, J. P. (2006). Is avoiding an aversive outcome rewarding? Neural substrates of avoidance learning in the human brain. *PLOS Biology*, 4, e233. <https://doi.org/10.1371/journal.pbio.0040233>
- Kim, J. J., & Jung, M. W. (2006). Neural circuits and mechanisms involved in Pavlovian fear conditioning: a critical review. *Neuroscience & Biobehavioral Reviews*, 30, 188–202. <https://doi.org/10.1016/j.neubiorev.2005.06.005>
- Kindt, M., & Soeter, M. (2014). Fear inhibition in high trait anxiety. *PLoS ONE*, 9. <https://doi.org/10.1371/journal.pone.0086462>
- Kishida, K. T., King-Casas, B., & Montague, P. R. (2010). Neuroeconomic approaches to mental disorders. *Neuron*, 67, 543–554. <https://doi.org/10.1016/j.neuron.2010.07.021>
- Kondo, N., Sembajwe, G., Kawachi, I., Dam, R. M. van, Subramanian, S. V., & Yamagata, Z. (2009). Income inequality, mortality, and self rated health: meta-analysis of multilevel studies. *BMJ*, 339, b4471. <https://doi.org/10.1136/bmj.b4471>

- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15, 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Kotov, R., Gamez, W., Schmidt, F., & Watson, D. (2010). Linking “big” personality traits to anxiety, depressive, and substance use disorders: a meta-analysis. *Psychological Bulletin*, 136, 768–821. <https://doi.org/10.1037/a0020327>
- Kotov, R., Krueger, R. F., Watson, D., Achenbach, T. M., Althoff, R. R., Bagby, R. M., ... Zimmerman, M. (2017). The hierarchical taxonomy of psychopathology (HiTOP): a dimensional alternative to traditional nosologies. *Journal of Abnormal Psychology*, 126, 454–477. <https://doi.org/10.1037/abn0000258>
- Kozak, M. J., & Miller, G. A. (1982). Hypothetical constructs versus intervening variables: a re-appraisal of the three-systems model of anxiety assessment. *Behavioral Assessment*, 4, 347–358.
- Kramer, M. D., Patrick, C. J., Krueger, R. F., & Gasperi, M. (2012). Delineating physiologic defensive reactivity in the domain of self-report: phenotypic and etiologic structure of dispositional fear. *Psychological Medicine*, 42, 1305–1320. <https://doi.org/10.1017/S0033291711002194>
- Kramer, M. D., Arbisi, P. A., Thuras, P. D., Krueger, R. F., Erbes, C. R., & Polusny, M. A. (2016). The class-dimensional structure of PTSD before and after deployment to Iraq: evidence from direct comparison of dimensional, categorical, and hybrid models. *Journal of Anxiety Disorders*, 39, 1–9. <https://doi.org/10.1016/j.janxdis.2016.02.004>
- Kramer, M. D., Patrick, C. J., Hettema, J. M., Moore, A. A., Sawyers, C. K., & Yancey, J. R. (2019). Quantifying dispositional fear as threat sensitivity: development and initial validation of a model-based scale measure. *Assessment*, 1073191119837613. <https://doi.org/10.1177/1073191119837613>
- Kristjansson, S. D., Kircher, J. C., & Webb, A. K. (2007). Multilevel models for repeated measures research designs in psychophysiology: an introduction to growth curve modeling. *Psychophysiology*, 44, 728–736. <https://doi.org/10.1111/j.1469-8986.2007.00544.x>
- Krueger, C., & Tian, L. (2004). A comparison of the general linear mixed model and repeated measures ANOVA using a dataset with multiple missing data points. *Biological Research for Nursing*, 6, 151–157.
- Krueger, R. F. (1999). The structure of common mental disorders. *Archives of General Psychiatry*, 56, 921–926. <https://doi.org/10.1001/archpsyc.56.10.921>
- Krueger, R. F., Derringer, J., Markon, K. E., Watson, D., & Skodol, A. E. (2012). Initial construction of a maladaptive personality trait model and inventory for DSM-5. *Psychological Medicine*, 42, 1879–1890. <https://doi.org/10.1017/S0033291711002674>
- Krueger, R. F., & DeYoung, C. G. (2016). The RDoC initiative and the structure of psychopathology. *Psychophysiology*, 53, 351–354. <https://doi.org/10.1111/psyp.12551>
- Krueger, R. F., Hopwood, C. J., Wright, A. G. C., & Markon, K. E. (2014). Challenges and strategies in helping the DSM become more dimensional and empirically based. *Current Psychiatry Reports*, 16, 1–6. <https://doi.org/10.1007/s11920-014-0515-3>

- Krueger, R. F., & Markon, K. E. (2006). Reinterpreting comorbidity: a model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology*, 2, 111–133. <https://doi.org/10.1146/annurev.clinpsy.2.022305.095213>
- Krueger, R. F., & Markon, K. E. (2014). The role of the DSM-5 personality trait model in moving toward a quantitative and empirically based approach to classifying personality and psychopathology. *Annual Review of Clinical Psychology*, 10, 477–501. <https://doi.org/10.1146/annurev-clinPSY-032813-153732>
- Krueger, R. F., McGue, M., & Iacono, W. G. (2001). The higher-order structure of common DSM mental disorders: internalization, externalization, and their connections to personality. *Personality and Individual Differences*, 30, 1245–1259. [https://doi.org/10.1016/S0191-8869\(00\)00106-9](https://doi.org/10.1016/S0191-8869(00)00106-9)
- Krueger, R. F., & Tackett, J. L. (2003). Personality and psychopathology: working toward the bigger picture. *Journal of Personality Disorders*, 17, 109–128. <https://doi.org/10.1521/pedi.17.2.109.23986>
- Krypotos, A.-M. (2015). Avoidance learning: a review of theoretical models and recent developments. *Frontiers in Behavioral Neuroscience*, 9. <https://doi.org/10.3389/fnbeh.2015.00189>
- Kudielka, B. M., & Kirschbaum, C. (2005). Sex differences in HPA axis responses to stress: a review. *Biological Psychology*, 69, 113–132.
- Kushner, S. C., Quilty, L. C., Tackett, J. L., & Bagby, R. M. (2011). The hierarchical structure of the dimensional assessment of personality pathology (DAPP-BQ). *Journal of Personality Disorders*, 25, 504–516. <https://doi.org/10.1521/pedi.2011.25.4.504>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). LmerTest package: tests in linear mixed effects models. *Journal of Statistical Software*, 82.
- Kvaal, K., Ulstein, I., Nordhus, I. H., & Engedal, K. (2005). The Spielberger state-trait anxiety inventory (STAI): the state scale in detecting mental disorders in geriatric patients. *International Journal of Geriatric Psychiatry*, 20, 629–634. <https://doi.org/10.1002/gps.1330>
- Lamers, S. M. A., Westerhof, G. J., Kovács, V., & Bohlmeijer, E. T. (2012). Differential relationships in the association of the Big Five personality traits with positive mental health and psychopathology. *Journal of Research in Personality*, 46, 517–524. <https://doi.org/10.1016/j.jrp.2012.05.012>
- Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1998). Emotion, motivation, and anxiety: brain mechanisms and psychophysiology. *Biological Psychiatry*, 44, 1248–1263.
- Lang, P. J., Levin, D. N., Miller, G. A., & Kozak, M. J. (1983). Fear behavior, fear imagery, and the psychophysiology of emotion: the problem of affective response integration. *Journal of Abnormal Psychology*, 92, 276–306. <https://doi.org/10.1037/0021-843X.92.3.276>
- Lang, P. J., McTeague, L. M., & Bradley, M. M. (2014). Pathological anxiety and function/dysfunction in the brain's fear/defense circuitry. *Restorative Neurology and Neuroscience*, 32, 63–77. <https://doi.org/10.3233/RNN-139012>
- Lange, I., Goossens, L., Bakker, J., Michielse, S., Marcelis, M., Wichers, M., ... Schruers, K. (2019). Functional neuroimaging of associative learning and

- generalization in specific phobia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 89, 275–285. <https://doi.org/10.1016/j.pnpbp.2018.09.008>
- Lanius, R. A., Bluhm, R., Lanius, U., & Pain, C. (2006). A review of neuroimaging studies in PTSD: heterogeneity of response to symptom provocation. *Journal of Psychiatric Research*, 40, 709–729. <https://doi.org/10.1016/j.jpsychires.2005.07.007>
- Larson, C. L., Ruffalo, D., Nietert, J. Y., & Davidson, R. J. (2000). Temporal stability of the emotion-modulated startle response. *Psychophysiology*, 37, 92–101.
- Lattin, J. M., Carroll, J. D., & Green, P. E. (2003). *Analyzing Multivariate Data*. Thomson Brooks/Cole.
- Latzman, R. D., & Masuda, A. (2013). Examining mindfulness and psychological inflexibility within the framework of Big Five personality. *Personality and Individual Differences*, 55, 129–134. <https://doi.org/10.1016/j.paid.2013.02.019>
- Lauriola, M., & Levin, I. P. (2001). Personality traits and risky decision-making in a controlled experimental task: an exploratory study. *Personality and Individual Differences*, 31, 215–226. [https://doi.org/10.1016/S0191-8869\(00\)00130-6](https://doi.org/10.1016/S0191-8869(00)00130-6)
- Lawrence, M. A. (2013). EZ: easy analysis and visualization of factorial experiments. *R Package Version*, 3.
- Lebowitz, E. R., Shic, F., Campbell, D., Basile, K., & Silverman, W. K. (2015). Anxiety sensitivity moderates behavioral avoidance in anxious youth. *Behaviour Research and Therapy*, 74, 11–17. <https://doi.org/10.1016/j.brat.2015.08.009>
- LeDoux, J. E., Moscarello, J., Sears, R., & Campese, V. (2017). The birth, death and resurrection of avoidance: a reconceptualization of a troubled paradigm. *Molecular Psychiatry*, 22, 24–36. <https://doi.org/10.1038/mp.2016.166>
- LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, 23, 155–184. <https://doi.org/10.1146/annurev.neuro.23.1.155>
- LeDoux, J. E. (2017). Semantics, surplus meaning, and the science of fear. *Trends in Cognitive Sciences*, 21, 303–306. <https://doi.org/10.1016/j.tics.2017.02.004>
- Lee, K., & Ashton, M. C. (2004). Psychometric properties of the HEXACO personality inventory. *Multivariate Behavioral Research*, 39, 329–358. https://doi.org/10.1207/s15327906mbr3902_8
- Lee, K., Ogunfowora, B., & Ashton, M. C. (2005). Personality traits beyond the Big Five: are they within the HEXACO space? *Journal of Personality*, 73, 1437–1463. <https://doi.org/10.1111/j.1467-6494.2005.00354.x>
- Lenaert, B., van de Ven, V., Kaas, A. L., & Vlaeyen, J. W. S. (2016). Generalization on the basis of prior experience is predicted by individual differences in working memory. *Behavior Therapy*, 47, 130–140. <https://doi.org/10.1016/j.beth.2015.10.001>
- Leung, P. W. L., & Wong, M. M. T. (1998). Can cognitive distortions differentiate between internalising and externalising problems? *The Journal of Child Psychology and Psychiatry and Allied Disciplines*, 39, 263–269. <https://doi.org/10.1017/S0021963097001868>
- Levy, D. J., & Glimcher, P. W. (2012). The root of all value: a neural common currency for choice. *Current Opinion in Neurobiology*, 22, 1027–1038. <https://doi.org/10.1016/j.conb.2012.06.001>

- Lewis, A. H., Niznikiewicz, M. A., Delamater, A. R., & Delgado, M. R. (2013). Avoidance-based human Pavlovian-to-instrumental transfer. *European Journal of Neuroscience*, 38, 3740–3748. <https://doi.org/10.1111/ejn.12377>
- Lieberman, L., Stevens, E. S., Funkhouser, C. J., Weinberg, A., Sarapas, C., Huggins, A. A., & Shankman, S. A. (2017). How many blinks are necessary for a reliable startle response? A test using the NPU-threat task. *International Journal of Psychophysiology*, 114, 24–30. <https://doi.org/10.1016/j.ijpsycho.2017.01.012>
- Lilienfeld, S. O. (1997). The relation of anxiety sensitivity to higher and lower order personality dimensions: implications for the etiology of panic attacks. *Journal of Abnormal Psychology*, 106, 539.
- Lilienfeld, S. O. (2014). The Research Domain Criteria (RDoC): an analysis of methodological and conceptual challenges. *Behaviour Research and Therapy*, 62, 129–139. <https://doi.org/10.1016/j.brat.2014.07.019>
- Lilienfeld, S. O., Jacob, R. G., & Turner, S. M. (1989). Comment on Holloway and McNally's (1987) "Effects of anxiety sensitivity on the response to hyperventilation." *Journal of Abnormal Psychology*, 98, 100–102.
- Lilienfeld, S. O., & Treadway, M. T. (2016). Clashing diagnostic approaches: DSM-ICD versus RDoC. *Annual Review of Clinical Psychology*, 12, 435–463. <https://doi.org/10.1146/annurev-clinPSY-021815-093122>
- Lilienfeld, S. O., Turner, S. M., & Jacob, R. G. (1993). Anxiety sensitivity: an examination of theoretical and methodological issues. *Advances in Behaviour Research and Therapy*, 15, 147–183. [https://doi.org/10.1016/0146-6402\(93\)90019-X](https://doi.org/10.1016/0146-6402(93)90019-X)
- Lissek, S. (2012). Toward an account of clinical anxiety predicated on basic, neurally mapped mechanisms of Pavlovian fear-learning: the case for conditioned overgeneralization. *Depression and Anxiety*, 29, 257–263. <https://doi.org/10.1002/da.21922>
- Lissek, S., Biggs, A. L., Rabin, S. J., Cornwell, B. R., Alvarez, R. P., Pine, D. S., & Grillon, C. (2008). Generalization of conditioned fear-potentiated startle in humans: experimental validation and clinical relevance. *Behaviour Research and Therapy*, 46, 678–687. <https://doi.org/10.1016/j.brat.2008.02.005>
- Lissek, S., Bradford, D. E., Alvarez, R. P., Burton, P., Espensen-Sturges, T., Reynolds, R. C., & Grillon, C. (2013). Neural substrates of classically conditioned fear-generalization in humans: a parametric fMRI study. *Social Cognitive and Affective Neuroscience*. <https://doi.org/10.1093/scan/nst096>
- Lissek, S., & Grillon, C. (2012). Learning models of PTSD. *The Oxford Handbook of Traumatic Stress Disorders*. <https://doi.org/10.1093/oxfordhb/9780195399066.013.0013>
- Lissek, S., Kaczkurkin, A. N., Rabin, S., Geraci, M., Pine, D. S., & Grillon, C. (2014). Generalized anxiety disorder is associated with overgeneralization of classically conditioned fear. *Biological Psychiatry*, 75, 909–915. <https://doi.org/10.1016/j.biopsych.2013.07.025>
- Lissek, S., Levenson, J., Biggs, A. L., Johnson, L. L., Ameli, R., Pine, D. S., & Grillon, C. (2008). Elevated fear conditioning to socially relevant unconditioned stimuli in social anxiety disorder. *The American Journal of Psychiatry*, 165, 124–132. <https://doi.org/10.1176/appi.ajp.2007.06091513>

- Lissek, S., Orme, K., McDowell, D. J., Johnson, L. L., Luckenbaugh, D. A., Baas, J. M., ... Grillon, C. (2007). Emotion regulation and potentiated startle across affective picture and threat-of-shock paradigms. *Biological Psychology*, 76, 124–133. <https://doi.org/10.1016/j.biopsycho.2007.07.002>
- Lissek, S., Pine, D. S., & Grillon, C. (2006). The strong situation: a potential impediment to studying the psychobiology and pharmacology of anxiety disorders. *Biological Psychology*, 72, 265–270.
- Lissek, S., Powers, A. S., McClure, E. B., Phelps, E. A., Woldehawariat, G., Grillon, C., & Pine, D. S. (2005). Classical fear conditioning in the anxiety disorders: a meta-analysis. *Behaviour Research and Therapy*, 43, 1391–1424. <https://doi.org/10.1016/j.brat.2004.10.007>
- Lissek, S., Rabin, S., Heller, R. E., Lukenbaugh, D., Geraci, M., Pine, D. S., & Grillon, C. (2010). Overgeneralization of conditioned fear as a pathogenic marker of panic disorder. *American Journal of Psychiatry*, 167, 47–55. <https://doi.org/10.1176/appi.ajp.2009.09030410>
- Lissek, S., Rabin, S. J., McDowell, D. J., Dvir, S., Bradford, D. E., Geraci, M., ... Grillon, C. (2009). Impaired discriminative fear-conditioning resulting from elevated fear responding to learned safety cues among individuals with panic disorder. *Behaviour Research and Therapy*, 47, 111–118. <https://doi.org/10.1016/j.brat.2008.10.017>
- Lissek, S., & van Meurs, B. (2014). Learning models of PTSD: theoretical accounts and psychobiological evidence. *International Journal of Psychophysiology*. <https://doi.org/10.1016/j.ijpsycho.2014.11.006>
- Lommen, M. J. J., Duta, M., Vanbrabant, K., Jong, R. de, Juechems, K., & Ehlers, A. (2017). Training discrimination diminishes maladaptive avoidance of innocuous stimuli in a fear conditioning paradigm. *PLOS ONE*, 12, e0184485. <https://doi.org/10.1371/journal.pone.0184485>
- Lommen, M. J. J., Engelhard, I. M., & van den Hout, M. A. (2010). Neuroticism and avoidance of ambiguous stimuli: better safe than sorry? *Personality and Individual Differences*, 49, 1001–1006. <https://doi.org/10.1016/j.paid.2010.08.012>
- Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., ... Merz, C. J. (2017). Don't fear 'fear conditioning': methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neuroscience & Biobehavioral Reviews*, 77, 247–285. <https://doi.org/10.1016/j.neubiorev.2017.02.026>
- Lonsdorf, T. B., & Merz, C. J. (2017). More than just noise: inter-individual differences in fear acquisition, extinction and return of fear in humans - biological, experiential, temperamental factors, and methodological pitfalls. *Neuroscience & Biobehavioral Reviews*, 80, 703–728. <https://doi.org/10.1016/j.neubiorev.2017.07.007>
- Lovibond, P. F., Saunders, J. C., Weidemann, G., & Mitchell, C. J. (2008). Evidence for expectancy as a mediator of avoidance and anxiety in a laboratory model of human avoidance learning. *The Quarterly Journal of Experimental Psychology*, 61, 1199–1216. <https://doi.org/10.1080/17470210701503229>
- Lüdtke, D. (2018). Sjstats: statistical functions for regression models. *R Package Version 0.14*, 3.

- Ludvik, D., Boschen, M. J., & Neumann, D. L. (2015). Effective behavioural strategies for reducing disgust in contamination-related OCD: a review. *Clinical Psychology Review*, 42, 116–129. <https://doi.org/10.1016/j.cpr.2015.07.001>
- Luyten, L., Vansteenwegen, D., van Kuyck, K., Gabriëls, L., & Nuttin, B. (2011). Contextual conditioning in rats as an animal model for generalized anxiety disorder. *Cognitive, Affective, & Behavioral Neuroscience*, 11, 228–244. <https://doi.org/10.3758/s13415-011-0021-6>
- Ly, V., & Roelofs, K. (2009). Social anxiety and cognitive expectancy of aversive outcome in avoidance conditioning. *Behaviour Research and Therapy*, 47, 840–847. <https://doi.org/10.1016/j.brat.2009.06.015>
- Lykken, D. T. (1968). Statistical significance in psychological research. *Psychological Bulletin*, 70, 151.
- Lykken, D. T. (1991). What's wrong with psychology anyway. In D. Cicchetti & W. M. Grove (Eds.), *Thinking Clearly About Psychology* (Vol. 1, pp. 3–39). University of Minnesota Press. Retrieved from https://books.google.com/books?hl=en&lr=&id=-CoSzt8JNAC&oi=fnd&pg=PA3&dq=Thinking+Clearly+about+Psychology,&ots=EVjNsth1T9&sig=xx-uFY_kebwaqn6YAels16vOxCA
- Maas, C. J. M., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling, 1, *Methodology*, 1, 86–92.
- MacCallum, R. C., Zhang, S., Preacher, K. J., & Rucker, D. D. (2002). On the practice of dichotomization of quantitative variables. *Psychological Methods*, 7, 19–40. <https://doi.org/10.1037//1082-989X.7.1.19>
- MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding and suppression effect. *Prevention Science : The Official Journal of the Society for Prevention Research*, 1, 173.
- Mackintosh, N. J. (1974). *The psychology of animal learning*. Academic Press.
- Mahaffey, B. L., Watson, D., Clark, L. A., & Kotov, R. (2016). Clinical and personality traits in emotional disorders: evidence of a common framework. *Journal of Abnormal Psychology*, 125, 758–767. <https://doi.org/10.1037/abn0000187>
- Mahan, A. L., & Ressler, K. J. (2012). Fear conditioning, synaptic plasticity and the amygdala: implications for posttraumatic stress disorder. *Trends in Neurosciences*, 35, 24–35. <https://doi.org/10.1016/j.tins.2011.06.007>
- Malouff, J. M., Thorsteinsson, E. B., & Schutte, N. S. (2005). The relationship between the five-factor model of personality and symptoms of clinical disorders: a meta-analysis. *Journal of Psychopathology and Behavioral Assessment*, 27, 101–114. <https://doi.org/10.1007/s10862-005-5384-y>
- Maren, S. (2001). Neurobiology of Pavlovian fear conditioning. *Annual Review of Neuroscience*, 24, 897–931. <http://dx.doi.org/10.1146/annurev.neuro.24.1.897>
- Markon, K. E., Krueger, R. F., & Watson, D. (2005). Delineating the structure of normal and abnormal personality: an integrative hierarchical approach. *Journal of Personality and Social Psychology*, 88, 139–157. <https://doi.org/10.1037/0022-3514.88.1.139>
- Marshall, R. D., Spitzer, R., & Liebowitz, M. R. (1999). Review and critique of the new DSM-IV diagnosis of acute stress disorder. *American Journal of Psychiatry*, 156, 1677–1685. <https://doi.org/10.1176/ajp.156.11.1677>

- Martínez, K. G., Castro-Couch, M., Franco-Chaves, J. A., Ojeda-Arce, B., Segura, G., Milad, M. R., & Quirk, G. J. (2012). Correlations between psychological tests and physiological responses during fear conditioning and renewal. *Biology of Mood & Anxiety Disorders*, 2, 16. <https://doi.org/10.1186/2045-5380-2-16>
- Mathews, A., Mackintosh, B., & Fulcher, E. P. (1997). Cognitive biases in anxiety and attention to threat. *Trends in Cognitive Sciences*, 1, 340–345.
- Mauro, R. (1990). Understanding love (left out variables error): a method for estimating the effects of omitted variables. *Psychological Bulletin*, 108, 314.
- McCoach, D. B. (2018). Multilevel modeling. *The Reviewer's Guide to Quantitative Methods in the Social Sciences*, 292–312.
- McCrae, R. R., & Costa, P. T. (1987). Validation of the five-factor model of personality across instruments and observers. *Journal of Personality and Social Psychology*, 52, 81.
- McCrae, R. R., Jang, K. L., Ando, J., Ono, Y., Yamagata, S., Riemann, R., ... Spinath, F. M. (2008). Substance and artifact in the higher-order factors of the Big Five. *Journal of Personality and Social Psychology*, 95, 442–455. <https://doi.org/10.1037/0022-3514.95.2.442>
- McCrae, R. R., & Löckenhoff, C. E. (2010). Self-regulation and the five-factor model of personality traits. In *Handbook of personality and self-regulation* (pp. 145–168). Wiley-Blackwell. <https://doi.org/10.1002/9781444318111.ch7>
- McFatter, R. M. (1994). Interactions in predicting mood from extraversion and neuroticism. *Journal of Personality and Social Psychology*, 9.
- McKay, D., Abramowitz, J. S., Calamari, J. E., Kyrios, M., Radomsky, A., Sookman, D., ... Wilhelm, S. (2004). A critical evaluation of obsessive-compulsive disorder subtypes: symptoms versus mechanisms. *Clinical Psychology Review*, 24, 283–313. <https://doi.org/10.1016/j.cpr.2004.04.003>
- McLean, C. P., & Anderson, E. R. (2009). Brave men and timid women? A review of the gender differences in fear and anxiety. *Clinical Psychology Review*, 29, 496–505. <https://doi.org/10.1016/j.cpr.2009.05.003>
- McNally, R. J. (1987). Preparedness and phobias: a review. *Psychological Bulletin*, 101, 283–303. <https://doi.org/10.1037/0033-2909.101.2.283>
- McNally, R. J. (1989). Is anxiety sensitivity distinguishable from trait anxiety? Reply to Lilienfeld, Jacob, and Turner (1989). *Journal of Abnormal Psychology*, 98, 193–194.
- McNally, R. J. (1990). Psychological approaches to panic disorder: a review. *Psychological Bulletin*, 108, 403. <https://doi.org/10.1037/0033-2909.108.3.403>
- McNally, R. J. (2002). Anxiety sensitivity and panic disorder. *Biological Psychiatry*, 52, 938–946. [https://doi.org/10.1016/S0006-3223\(02\)01475-0](https://doi.org/10.1016/S0006-3223(02)01475-0)
- McNally, R. J. (2003). Progress and controversy in the study of posttraumatic stress disorder. *Annual Review of Psychology*, 54, 229–252. <https://doi.org/10.1146/annurev.psych.54.101601.145112>
- McNally, R. J. (2007). Mechanisms of exposure therapy: how neuroscience can improve psychological treatments for anxiety disorders. *Clinical Psychology Review*, 27, 750–759. <https://doi.org/10.1016/j.cpr.2007.01.003>

- McNaughton, N. (2011). Trait anxiety, trait fear and emotionality: the perspective from non-human studies. *Personality and Individual Differences*, 50, 898–906. <https://doi.org/10.1016/j.paid.2010.07.011>
- McNaughton, N., DeYoung, C. G., & Corr, P. J. (2016). Approach/avoidance. In J. R. Absher & J. Cloutier (Eds.), *Neuroimaging Personality, Social Cognition, and Character* (pp. 25–49). San Diego: Academic Press. <https://doi.org/10.1016/B978-0-12-800935-2.00002-6>
- McNeish, D. M., & Stapleton, L. M. (2016). The effect of small sample size on two-level model estimates: a review and illustration. *Educational Psychology Review*, 28, 295–314. <https://doi.org/10.1007/s10648-014-9287-x>
- McNeish, D., & Stapleton, L. M. (2016). Modeling clustered data with very few clusters. *Multivariate Behavioral Research*, 51, 495–518. <https://doi.org/10.1080/00273171.2016.1167008>
- Meehl, P. E. (1972). Specific genetic etiology, psychodynamics, and therapeutic nihilism. *International Journal of Mental Health*, 1, 10–27.
- Meehl, P. E. (1995). Bootstraps taxometrics: solving the classification problem in psychopathology. *American Psychologist*, 50, 266–275. <https://doi.org/10.1037/0003-066X.50.4.266>
- Mehta, P. D., & Neale, M. C. (2005). People are variables too: multilevel structural equations modeling. *Psychological Methods*, 10, 259.
- Mertens, G., & De Houwer, J. (2016). Potentiation of the startle reflex is in line with contingency reversal instructions rather than the conditioning history. *Biological Psychology*, 113, 91–99. <https://doi.org/10.1016/j.biopsycho.2015.11.014>
- Meyer, R. D., Dalal, R. S., & Hermida, R. (2010). A review and synthesis of situational strength in the organizational sciences. *Journal of Management*, 36, 121–140.
- Mikulincer, M., & Nizan, B. (1988). Causal attribution, cognitive interference, and the generalization of learned helplessness. *Journal of Personality and Social Psychology*, 55, 470–478. <https://doi.org/10.1037/0022-3514.55.3.470>
- Milad, M. R., Furtak, S. C., Greenberg, J. L., Keshaviah, A., Im, J. J., Falkenstein, M. J., ... Wilhelm, S. (2013). Deficits in conditioned fear extinction in obsessive-compulsive disorder and neurobiological changes in the fear circuit. *JAMA Psychiatry*, 70, 608–618. <https://doi.org/10.1001/jamapsychiatry.2013.914>
- Milgram, S., & Gudehus, C. (1978). *Obedience to authority*. Ziff-Davis Publishing Company.
- Miller, M. W., Wolf, E. J., Reardon, A., Greene, A., Ofrat, S., & McInerney, S. (2012). Personality and the latent structure of PTSD comorbidity. *Journal of Anxiety Disorders*, 26, 599–607. <https://doi.org/10.1016/j.janxdis.2012.02.016>
- Mineka, S. (1979). The role of fear in theories of avoidance learning, flooding, and extinction. *Psychological Bulletin*, 86, 985–1010. <https://doi.org/10.1037/0033-2909.86.5.985>
- Mineka, S., & Zinbarg, R. (2006). A contemporary learning theory perspective on the etiology of anxiety disorders: it's not what you thought it was. *American Psychologist*, 61, 10–26. <https://doi.org/10.1037/0003-066X.61.1.10>
- Minnen, A. van, Harned, M. S., Zoellner, L., & Mills, K. (2012). Examining potential contraindications for prolonged exposure therapy for PTSD. *European Journal of Psychotraumatology*, 3, 18805. <https://doi.org/10.3402/ejpt.v3i0.18805>

- Misangyi, V. F., LePine, J. A., Algina, J., & Goeddeke, F. (2006). The adequacy of repeated-measures regression for multilevel research: comparisons with repeated-measures ANOVA, multivariate repeated-measures ANOVA, and multilevel modeling across various multilevel research designs. *Organizational Research Methods*, 9, 5–28. <https://doi.org/10.1177/1094428105283190>
- Mischel, W. (1977). The interaction of person and situation. In D. Magnusson & N. S. Endler (Eds.), *Personality at the crossroads: Current issues in interactional psychology* (pp. 333–352). Hillsdale, NJ: Lawrence Erlbaum.
- Mischel, W. (1999). Implications of person–situation interaction: getting over the field's borderline personality disorder. *European Journal of Personality*, 13, 455–461. [https://doi.org/10.1002/\(SICI\)1099-0984\(199909/10\)13:5<455::AID-PER357>3.0.CO;2-5](https://doi.org/10.1002/(SICI)1099-0984(199909/10)13:5<455::AID-PER357>3.0.CO;2-5)
- Mkrtchian, A., Aylward, J., Dayan, P., Roiser, J. P., & Robinson, O. J. (2017). Modeling avoidance in mood and anxiety disorders using reinforcement learning. *Biological Psychiatry*, 82, 532–539. <https://doi.org/10.1016/j.biopsych.2017.01.017>
- Moeller, J. (2015). A word on standardization in longitudinal studies: don't. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.01389>
- Morey, L. C., & Boggs, C. (1991). *Personality assessment inventory (PAI)*. PAR Lutz, FL.
- Morey, L. C., Gunderson, J. G., Quigley, B. D., Shea, M. T., Skodol, A. E., McGlashan, T. H., ... Zanarini, M. C. (2002). The representation of borderline, avoidant, obsessive-compulsive, and schizotypal personality disorders by the five-factor model. *Journal of Personality Disorders*, 16, 215–234.
- Morey, L. C., Gunderson, J., Quigley, B. D., & Lyons, M. (2000). Dimensions and categories: the "Big Five" factors and the DSM personality disorders. *Assessment*, 7, 203–216.
- Morey, R. A., Dunsmoor, J. E., Haswell, C. C., Brown, V. M., Vora, A., Weiner, J., ... LaBar, K. S. (2015). Fear learning circuitry is biased toward generalization of fear associations in posttraumatic stress disorder. *Translational Psychiatry*, 5, e700. <https://doi.org/10.1038/tp.2015.196>
- Morriss, J. (2019). What do I do now? Intolerance of uncertainty is associated with discrete patterns of anticipatory physiological responding to different uncertain contexts. *BioRxiv*, 527317. <https://doi.org/10.1101/527317>
- Morriss, J., Christakou, A., & Reekum, C. M. van. (2015). Intolerance of uncertainty predicts fear extinction in amygdala-ventromedial prefrontal cortical circuitry. *Biology of Mood & Anxiety Disorders*, 5, 4. <https://doi.org/10.1186/s13587-015-0019-8>
- Morriss, J., Christakou, A., & van Reekum, C. M. (2016). Nothing is safe: intolerance of uncertainty is associated with compromised fear extinction learning. *Biological Psychology*, 121, Part B, 187–193. <https://doi.org/10.1016/j.biopsycho.2016.05.001>
- Morriss, J., Saldarini, F., Chapman, C., Pollard, M., & van Reekum, C. M. (2019). Out with the old and in with the new: the role of intolerance of uncertainty in reversal of threat and safety. *Journal of Experimental Psychopathology*, 10, 2043808719834451. <https://doi.org/10.1177/2043808719834451>

- Morriss, J., Saldarini, F., & Reekum, C. M. van. (2018). The devil is in the detail: the role of threat level and intolerance of uncertainty in extinction. *BioRxiv*, 479543. <https://doi.org/10.1101/479543>
- Morriss, J., & van Reekum, C. M. (2019). I feel safe when I know: contingency instruction promotes threat extinction in high intolerance of uncertainty individuals. *Behaviour Research and Therapy*, 116, 111–118. <https://doi.org/10.1016/j.brat.2019.03.004>
- Moscovitch, D. A. (2009). What is the core fear in social phobia? A new model to facilitate individualized case conceptualization and treatment. *Cognitive and Behavioral Practice*, 16, 123–134. <https://doi.org/10.1016/j.cbpra.2008.04.002>
- Mowrer, O. H. (1951). Two-factor learning theory: summary and comment. *Psychological Review*, 58, 350–354. <https://doi.org/10.1037/h0058956>
- Musca, S. C., Kamiejski, R., Nugier, A., Méot, A., Er-Rafiy, A., & Brauer, M. (2011). Data with hierarchical structure: impact of intraclass correlation and sample size on type-i error. *Frontiers in Psychology*, 2. <https://doi.org/10.3389/fpsyg.2011.00074>
- Mykletun, A., Bjerkeset, O., Øverland, S., Prince, M., Dewey, M., & Stewart, R. (2009). Levels of anxiety and depression as predictors of mortality: the hunt study. *The British Journal of Psychiatry*, 195, 118–125. <https://doi.org/10.1192/bjp.bp.108.054866>
- Mykletun, A., Overland, S., Aarø, L. E., Liabø, H.-M., & Stewart, R. (2008). Smoking in relation to anxiety and depression: evidence from a large population survey: the hunt study. *European Psychiatry*, 23, 77–84. <https://doi.org/10.1016/j.eurpsy.2007.10.005>
- Najavits, L. M. (2015). The problem of dropout from “gold standard” PTSD therapies. *F1000Prime Reports*, 7. <https://doi.org/10.12703/P7-43>
- Nakagawa, S., Johnson, P. C., & Schielzeth, H. (2017). The coefficient of determination R^2 and intra-class correlation coefficient from generalized linear mixed-effects models revisited and expanded. *Journal of the Royal Society Interface*, 14, 20170213.
- Nakagawa, S., & Schielzeth, H. (2013). A general and simple method for obtaining R^2 from generalized linear mixed-effects models. *Methods in Ecology and Evolution*, 4, 133–142. <https://doi.org/10.1111/j.2041-210x.2012.00261.x>
- Nanbu, M., Kurayama, T., Nakazawa, K., Matsuzawa, D., Komiya, Z., Haraguchi, T., ... Shimizu, E. (2010). Impaired P50 suppression in fear extinction in obsessive-compulsive disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 34, 317–322. <https://doi.org/10.1016/j.pnpbp.2009.12.005>
- Naragon-Gainey, K. (2010). Meta-analysis of the relations of anxiety sensitivity to the depressive and anxiety disorders. *Psychological Bulletin*, 136, 128–150. <https://doi.org/10.1037/a0018055>
- Naragon-Gainey, K., & Simms, L. J. (2017). Three-way interaction of neuroticism, extraversion, and conscientiousness in the internalizing disorders: evidence of disorder specificity in a psychiatric sample. *Journal of Research in Personality*. <https://doi.org/10.1016/j.jrp.2017.05.003>
- Naragon-Gainey, K., Watson, D., & Markon, K. E. (2009). Differential relations of depression and social anxiety symptoms to the facets of extraversion/positive

- emotionality. *Journal of Abnormal Psychology*, 118, 299–310.
<https://doi.org/10.1037/a0015637>
- Nash, J. C. (2014). On best practice optimization methods in R. *Journal of Statistical Software*, 60, 1–14.
- National Institute of Mental Health. (2016). RDoC matrix. Retrieved May 30, 2019, from <https://www.nimh.nih.gov/research/research-funded-by-nimh/RDoC/constructs/RDoC-matrix.shtml>
- Nawijn, L., van Zuiden, M., Frijling, J. L., Koch, S. B. J., Veltman, D. J., & Olff, M. (2015). Reward functioning in PTSD: a systematic review exploring the mechanisms underlying anhedonia. *Neuroscience & Biobehavioral Reviews*, 51, 189–204. <https://doi.org/10.1016/j.neubiorev.2015.01.019>
- Nelson, B. D., & Shankman, S. A. (2011). Does intolerance of uncertainty predict anticipatory startle responses to uncertain threat? *International Journal of Psychophysiology*, 81, 107–115. <https://doi.org/10.1016/j.ijpsycho.2011.05.003>
- Nelson, B. D., Weinberg, A., Pawluk, J., Gawlowska, M., & Proudfit, G. H. (2014). An event-related potential investigation of fear generalization and intolerance of uncertainty. *Behavior Therapy*. <https://doi.org/10.1016/j.beth.2014.09.010>
- Nolen-Hoeksema, S. (2012). Emotion regulation and psychopathology: the role of gender. *Annual Review of Clinical Psychology*, 8, 161–187.
<https://doi.org/10.1146/annurev-clinpsy-032511-143109>
- Ocañez, K. L. S., McHugh, R. K., & Otto, M. W. (2010). A meta-analytic review of the association between anxiety sensitivity and pain. *Depression and Anxiety*, 27, 760–767. <https://doi.org/10.1002/da.20681>
- Olatunji, B. O., Cisler, J. M., Meunier, S., Connolly, K., & Lohr, J. M. (2008). Expectancy bias for fear and disgust and behavioral avoidance in spider fearful individuals. *Cognitive Therapy and Research*, 32, 460–469.
<https://doi.org/10.1007/s10608-007-9164-8>
- Olatunji, B. O., Moretz, M. W., & Zlomke, K. R. (2010). Linking cognitive avoidance and GAD symptoms: the mediating role of fear of emotion. *Behaviour Research and Therapy*, 48, 435–441. <https://doi.org/10.1016/j.brat.2009.11.014>
- Olatunji, B. O., & Wolitzky-Taylor, K. B. (2009). Anxiety sensitivity and the anxiety disorders: a meta-analytic review and synthesis. *Psychological Bulletin*, 135, 974.
- Oltmanns, J. R., & Widiger, T. A. (2018). Assessment of fluctuation between grandiose and vulnerable narcissism: development and initial validation of the flux scales. *Psychological Assessment*, 13.
- Ong, C. W., Lee, E. B., & Twohig, M. P. (2018). A meta-analysis of dropout rates in acceptance and commitment therapy. *Behaviour Research and Therapy*.
<https://doi.org/10.1016/j.brat.2018.02.004>
- Ormel, J., Jeronimus, B. F., Kotov, R., Riese, H., Bos, E. H., Hankin, B., ... Oldehinkel, A. J. (2013). Neuroticism and common mental disorders: meaning and utility of a complex relationship. *Clinical Psychology Review*, 33, 686–697.
<https://doi.org/10.1016/j.cpr.2013.04.003>
- Öst, L.-G. (1989). One-session treatment for specific phobias. *Behaviour Research and Therapy*, 27, 1–7.
- Öst, L.-G. (1996). One-session group treatment of spider phobia. *Behaviour Research and Therapy*, 34, 707–715.

- Otto, M. W., Leyro, T. M., Christian, K., Deveney, C. M., Reese, H., Pollack, M. H., & Orr, S. P. (2007). Prediction of “fear” acquisition in healthy control participants in a de novo fear-conditioning paradigm. *Behavior Modification*, 31, 32–51. <https://doi.org/10.1177/0145445506295054>
- Ozomaro, U., Wahlestedt, C., & Nemeroff, C. B. (2013). Personalized medicine in psychiatry: problems and promises. *BMC Medicine*, 11, 132. <https://doi.org/10.1186/1741-7015-11-132>
- Panitz, C., Sperl, M. F. J., Hennig, J., Klucken, T., Hermann, C., & Mueller, E. M. (2018). Fearfulness, neuroticism/anxiety, and COMT Val158Met in long-term fear conditioning and extinction. *Neurobiology of Learning and Memory*. <https://doi.org/10.1016/j.nlm.2018.06.001>
- Patrick, C. J., & Hajcak, G. (2016). RDoC: translating promise into progress. *Psychophysiology*, 53, 415–424. <https://doi.org/10.1111/psyp.12612>
- Patrick, C. J., Venables, N. C., Yancey, J. R., Hicks, B. M., Nelson, L. D., & Kramer, M. D. (2013). A construct-network approach to bridging diagnostic and physiological domains: application to assessment of externalizing psychopathology. *Journal of Abnormal Psychology*, 122, 902.
- Paulhus, D. L., Robins, R. W., Trzesniewski, K. H., & Tracy, J. L. (2004). Two replicable suppressor situations in personality research. *Multivariate Behavioral Research*, 39, 303–328. https://doi.org/10.1207/s15327906mbr3902_7
- Paulus, M. P., & Stein, M. B. (2006). An insular view of anxiety. *Biological Psychiatry*, 60, 383–387. <https://doi.org/10.1016/j.biopsych.2006.03.042>
- Paunonen, S. V. (1998). Hierarchical organization of personality and prediction of behavior. *Journal of Personality and Social Psychology*, 74, 538–556. <https://doi.org/10.1037/0022-3514.74.2.538>
- Paunonen, S. V., & Ashton, M. C. (2001). Big five factors and facets and the prediction of behavior. *Journal of Personality and Social Psychology*, 81, 524–539. <https://doi.org/10.1037/0022-3514.81.3.524>
- Pavlov, I. P. (1927). *Conditioned reflexes*. Courier Dover Publications.
- Peckham, A. D., & Johnson, S. L. (2015). Spontaneous eye-blink rate as an index of reward responsivity validation and links to bipolar disorder. *Clinical Psychological Science*. <https://doi.org/10.1177/2167702615594999>
- Perkins, A. M., Cooper, A., Abdelall, M., Smillie, L. D., & Corr, P. J. (2010). Personality and defensive reactions: fear, trait anxiety, and threat magnification. *Journal of Personality*, 78, 1071–1090. <https://doi.org/10.1111/j.1467-6494.2010.00643.x>
- Perusini, J. N., & Fanselow, M. S. (2015). Neurobehavioral perspectives on the distinction between fear and anxiety. *Learning & Memory*, 22, 417–425. <https://doi.org/10.1101/lm.039180.115>
- Peterson, R. A., & Heilbronner, R. L. (1987). The anxiety sensitivity index: construct validity and factor analytic structure. *Journal of Anxiety Disorders*, 1, 117–121. [https://doi.org/10.1016/0887-6185\(87\)90002-8](https://doi.org/10.1016/0887-6185(87)90002-8)
- Phelps, E. A., Delgado, M. R., Nearing, K. I., & LeDoux, J. E. (2004). Extinction learning in humans: role of the amygdala and vmPFC. *Neuron*, 43, 897–905. <https://doi.org/10.1016/j.neuron.2004.08.042>

- Pineles, S. L., Vogt, D. S., & Orr, S. P. (2009). Personality and fear responses during conditioning: beyond extraversion. *Personality and Individual Differences*, 46, 48–53. <https://doi.org/10.1016/j.paid.2008.09.003>
- Pitman, R. K. (1988). Post-traumatic stress disorder, conditioning, and network theory. *Psychiatric Annals*, 18, 182–189. <https://doi.org/10.3928/0048-5713-19880301-11>
- Pittenger, D. J. (2004). The limitations of extracting typologies from trait measures of personality. *Personality and Individual Differences*, 37, 779–787.
- Pittig, A., Alpers, G. W., Niles, A. N., & Craske, M. G. (2015). Avoidant decision-making in social anxiety disorder: a laboratory task linked to in vivo anxiety and treatment outcome. *Behaviour Research and Therapy*, 73, 96–103. <https://doi.org/10.1016/j.brat.2015.08.003>
- Pittig, A., Brand, M., Pawlikowski, M., & Alpers, G. W. (2014). The cost of fear: avoidant decision making in a spider gambling task. *Journal of Anxiety Disorders*, 28, 326–334. <https://doi.org/10.1016/j.janxdis.2014.03.001>
- Pittig, A., & Dehler, J. (2018). Same fear responses, less avoidance: rewards competing with aversive outcomes do not buffer fear acquisition, but attenuate avoidance to accelerate subsequent fear extinction. *Behaviour Research and Therapy*. <https://doi.org/10.1016/j.brat.2018.11.003>
- Pittig, A., Hengen, K., Bublatzky, F., & Alpers, G. W. (2018). Social and monetary incentives counteract fear-driven avoidance: evidence from approach-avoidance decisions. *Journal of Behavior Therapy and Experimental Psychiatry*, 60, 69–77. <https://doi.org/10.1016/j.jbtep.2018.04.002>
- Pittig, A., Treanor, M., LeBeau, R. T., & Craske, M. G. (2018). The role of associative fear and avoidance learning in anxiety disorders: gaps and directions for future research. *Neuroscience & Biobehavioral Reviews*, 88, 117–140. <https://doi.org/10.1016/j.neubiorev.2018.03.015>
- Poletti, S., Radaelli, D., Cucchi, M., Ricci, L., Vai, B., Smeraldi, E., & Benedetti, F. (2015). Neural correlates of anxiety sensitivity in panic disorder: a functional magnetic resonance imaging study. *Psychiatry Research: Neuroimaging*, 233, 95–101. <https://doi.org/10.1016/j.psychresns.2015.05.013>
- Preacher, K. J., Curran, P. J., & Bauer, D. J. (2006). Computational tools for probing interactions in multiple linear regression, multilevel modeling, and latent curve analysis. *Journal of Educational and Behavioral Statistics*, 31, 437–448. <https://doi.org/10.3102/10769986031004437>
- Preacher, K. J., Zyphur, M. J., & Zhang, Z. (2010). A general multilevel SEM framework for assessing multilevel mediation. *Psychological Methods*, 15, 209–233. <https://doi.org/10.1037/a0020141>
- Putwain, D. W., Symes, W., & Wilkinson, H. M. (2017). Fear appeals, engagement, and examination performance: the role of challenge and threat appraisals. *British Journal of Educational Psychology*, 87, 16–31. <https://doi.org/10.1111/bjep.12132>
- Quilty, L. C., Ayearst, L., Chmielewski, M., Pollock, B. G., & Bagby, R. M. (2013). The psychometric properties of the personality inventory for DSM-5 in an APA DSM-5 field trial sample. *Assessment*, 20, 362–369. <https://doi.org/10.1177/1073191113486183>

- Quilty, L. C., DeYoung, C. G., Oakman, J. M., & Bagby, R. M. (2014). Extraversion and behavioral activation: integrating the components of approach. *Journal of Personality Assessment*, 96, 87–94.
<https://doi.org/10.1080/00223891.2013.834440>
- Rabe-Hesketh, S., Skrondal, A., & Pickles, A. (2004). Generalized multilevel structural equation modeling. *Psychometrika*, 69, 167–190.
- Rachman, S. (1991). Neo-conditioning and the classical theory of fear acquisition. *Clinical Psychology Review*, 11, 155–173. [https://doi.org/10.1016/0272-7358\(91\)90093-A](https://doi.org/10.1016/0272-7358(91)90093-A)
- Rachman, S. (2015). The evolution of behaviour therapy and cognitive behaviour therapy. *Behaviour Research and Therapy*, 64, 1–8.
<https://doi.org/10.1016/j.brat.2014.10.006>
- Rachman, S. (2004). Fear and courage: a psychological perspective. *Social Research: An International Quarterly*, 71, 149–176.
- Radell, M. L., Beck, K. D., Gilbertson, M. W., & Myers, C. E. (2017). Post-traumatic stress disorder symptom burden and gender each affect generalization in a reward- and punishment-learning task. *PLOS ONE*, 12, e0172144.
<https://doi.org/10.1371/journal.pone.0172144>
- Rapee, R. M., & Heimberg, R. G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy*, 35, 741–756.
- Rapee, R. M., & Spence, S. H. (2004). The etiology of social phobia: empirical evidence and an initial model. *Clinical Psychology Review*, 24, 737–767.
<https://doi.org/10.1016/j.cpr.2004.06.004>
- Ratcliff, R., Philastides, M. G., & Sajda, P. (2009). Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. *Proceedings of the National Academy of Sciences*, 106, 6539–6544.
- Rauch, S. A., Eftekhari, A., & Ruzek, J. I. (2012). Review of exposure therapy: a gold standard for PTSD treatment. *Journal of Rehabilitation Research and Development*, 49, 679–688.
- Rauch, S. L., Milad, M. R., Orr, S. P., Quinn, B. T., Fischl, B., & Pitman, R. K. (2005). Orbitofrontal thickness, retention of fear extinction, and extraversion. *Neuroreport*, 16, 1909–1912.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). Sage.
- RCore Team. (2014). R: a language and environment for statistical computing: <http://www.R-Project.org>.
- Rector, N. A., Hood, K., Richter, M. A., & Michael Bagby, R. (2002). Obsessive-compulsive disorder and the five-factor model of personality: distinction and overlap with major depressive disorder. *Behaviour Research and Therapy*, 40, 1205–1219. [https://doi.org/10.1016/S0005-7967\(02\)00024-4](https://doi.org/10.1016/S0005-7967(02)00024-4)
- Rector, N. A., Bagby, R. M., Huta, V., & Ayearst, L. E. (2012). Examination of the trait facets of the five-factor model in discriminating specific mood and anxiety disorders. *Psychiatry Research*, 199, 131–139.
<https://doi.org/10.1016/j.psychres.2012.04.027>
- Ree, M. J., French, D., MacLeod, C., & Locke, V. (2008). Distinguishing cognitive and somatic dimensions of state and trait anxiety: development and validation of the

- state-trait inventory for cognitive and somatic anxiety (STICSA). *Behavioural and Cognitive Psychotherapy*, 36, 313–332.
<http://dx.doi.org/10.1017/S1352465808004232>
- Reger, G. M., Hoffman, J., Riggs, D., Rothbaum, B. O., Ruzek, J., Holloway, K. M., & Kuhn, E. (2013). The “PE Coach” smartphone application: an innovative approach to improving implementation, fidelity, and homework adherence during prolonged exposure. *Psychological Services*, 10, 342–349.
<https://doi.org/10.1037/a0032774>
- Reiss, S. (1997). Trait anxiety: it’s not what you think it is. *Journal of Anxiety Disorders*, 11, 201–214. [https://doi.org/10.1016/S0887-6185\(97\)00006-6](https://doi.org/10.1016/S0887-6185(97)00006-6)
- Reiss, S., Peterson, R. A., Gursky, D. M., & McNally, R. J. (1986). Anxiety sensitivity, anxiety frequency and the prediction of fearfulness. *Behaviour Research and Therapy*, 24, 1–8. [https://doi.org/10.1016/0005-7967\(86\)90143-9](https://doi.org/10.1016/0005-7967(86)90143-9)
- Rescorla, R. A., & Solomon, R. L. (1967). Two-process learning theory: relationships between Pavlovian conditioning and instrumental learning. *Psychological Review*, 74, 151.
- Rescorla, R. A., & Wagner, A. (1972). A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. In A. Black & W. Prokasy (Eds.), *Classical Conditioning II: Current Research and Theory* (pp. 64–99). Appleton-Century-Crofts.
- Reyna, V. F., & Huettel, S. A. (2014). Reward, representation, and impulsivity: a theoretical framework for the neuroscience of risky decision making. In *The neuroscience of risky decision making* (pp. 11–42). Washington, DC, US: American Psychological Association. <https://doi.org/10.1037/14322-002>
- Rinck, M., Koene, M., Telli, S., Brink, W. M. den, Verhoeven, B., & Becker, E. S. (2016). The time course of location-avoidance learning in fear of spiders. *Cognition and Emotion*, 30, 430–443.
<https://doi.org/10.1080/02699931.2015.1009873>
- Roberts, B. W., Chernyshenko, O. S., Stark, S., & Goldberg, L. R. (2005). The structure of conscientiousness: an empirical investigation based on seven major personality questionnaires. *Personnel Psychology*, 58, 103–139.
<https://doi.org/10.1111/j.1744-6570.2005.00301.x>
- Roberts, B. W., Jackson, J. J., Burger, J., & Trautwein, U. (2009). Conscientiousness and externalizing psychopathology: overlap, developmental patterns, and etiology of two related constructs. *Development and Psychopathology*, 21, 871–888.
<https://doi.org/10.1017/S0954579409000479>
- Roberts, B. W., Lejuez, C., Krueger, R. F., Richards, J. M., & Hill, P. L. (2014a). What is conscientiousness and how can it be assessed? *Developmental Psychology*, 50, 1315–1330. <https://doi.org/10.1037/a0031109>
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychological Bulletin*, 132, 1.
- Rosellini, A. J., & Brown, T. A. (2011). The NEO five-factor inventory: latent structure and relationships with dimensions of anxiety and depressive disorders in a large clinical sample. *Assessment*, 18, 27–38.
<https://doi.org/10.1177/1073191110382848>

- Rosenbaum, B. L., Bui, E., Marin, M.-F., Holt, D. J., Lasko, N. B., Pitman, R. K., ... Milad, M. R. (2015). Demographic factors predict magnitude of conditioned fear. *International Journal of Psychophysiology*.
<https://doi.org/10.1016/j.ijpsycho.2015.06.010>
- Rothbaum, B. O., Meadows, E. A., Resick, P., & Foy, D. W. (2000). Cognitive-behavioral therapy. In E. B. Foa, T. M. Keane, & M. J. Friedman (Eds.), *Effective treatments for PTSD: Practice guidelines from the International Society for Traumatic Stress Studies* (pp. 320–325). New York, NY, US: Guilford Press.
- Rouhani, N., Wimmer, G. E., Schneier, F. R., Fyer, A. J., Shohamy, D., & Simpson, H. B. (2018). Impaired generalization of reward but not loss in obsessive–compulsive disorder. *Depression and Anxiety*, 0. <https://doi.org/10.1002/da.22857>
- RStudio Team. (2015). RStudio: integrated development for r. RStudio, Inc., Boston, MA URL <http://www.Rstudio.com>, 42, 14.
- Ruggero, C. (2018, August 18). Integrating a dimensional, hierarchical taxonomy of psychopathology into clinical practice. doi:10.31234/osf.io/r2jt6
- Sachs, G., Anderer, P., Doby, D., Saletu, B., & Dantendorfer, K. (2003). Impaired conditional discrimination learning in social phobia. *Neuropsychobiology*, 47, 66–72. <https://doi.org/10.1159/000070011>
- Salkovskis, P. M., Clark, D. M., Hackmann, A., Wells, A., & Gelder, M. G. (1999). An experimental investigation of the role of safety-seeking behaviours in the maintenance of panic disorder with agoraphobia. *Behaviour Research and Therapy*, 37, 559–574. [https://doi.org/10.1016/S0005-7967\(98\)00153-3](https://doi.org/10.1016/S0005-7967(98)00153-3)
- Samuel, D. B., Simms, L. J., Clark, L. A., Livesley, W. J., & Widiger, T. A. (2010). An item response theory integration of normal and abnormal personality scales. *Personality Disorders*, 1, 5–21. <https://doi.org/10.1037/a0018136>
- Samuel, D. B., & Widiger, T. A. (2011). Conscientiousness and obsessive-compulsive personality disorder. *Personality Disorders: Theory, Research, and Treatment*, 2, 161.
- Sandin, B., Chorot, P., & McNally, R. J. (2001). Anxiety sensitivity index: normative data and its differentiation from trait anxiety. *Behaviour Research and Therapy*, 39, 213–219. [https://doi.org/10.1016/S0005-7967\(00\)00009-7](https://doi.org/10.1016/S0005-7967(00)00009-7)
- Sandt, A. R., Sloan, D. M., & Johnson, K. J. (2009). Measuring appetitive responding with the postauricular reflex. *Psychophysiology*, 46, 491–497.
<https://doi.org/10.1111/j.1469-8986.2009.00797.x>
- Sarinopoulos, I., Grupe, D. W., Mackiewicz, K. L., Herrington, J. D., Lor, M., Steege, E. E., & Nitschke, J. B. (2010). Uncertainty during anticipation modulates neural responses to aversion in human insula and amygdala. *Cerebral Cortex*, 20, 929–940. <https://doi.org/10.1093/cercor/bhp155>
- Sartorius, N., Üstün, T. B., Lecrubier, Y., & Wittchen, H.-U. (1996). Depression comorbid with anxiety: results from the who study on psychological disorders in primary health care. *The British Journal of Psychiatry*, 168, 38–43.
<https://doi.org/10.1192/S0007125000298395>
- Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin*, 2, 110–114. <https://doi.org/10.2307/3002019>
- Sauer-Zavala, S., Boswell, J. F., Gallagher, M. W., Bentley, K. H., Ametaj, A., & Barlow, D. H. (2012). The role of negative affectivity and negative reactivity to

- emotions in predicting outcomes in the unified protocol for the transdiagnostic treatment of emotional disorders. *Behaviour Research and Therapy*, 50, 551–557. <https://doi.org/10.1016/j.brat.2012.05.005>
- Schad, D. J., Garbusow, M., Friedel, E., Sommer, C., Sebold, M., Hägele, C., ... Rapp, M. A. (2019). Neural correlates of instrumental responding in the context of alcohol-related cues index disorder severity and relapse risk. *European Archives of Psychiatry and Clinical Neuroscience*, 269, 295–308. <https://doi.org/10.1007/s00406-017-0860-4>
- Schmidt, N. B. (1999). Examination of differential anxiety sensitivities in panic disorder: a test of anxiety sensitivity subdomains predicting fearful responding to a 35% CO₂ challenge. *Cognitive Therapy and Research*, 23, 3–20. <https://doi.org/10.1023/A:1018754522981>
- Schmidt, N. B., & Joiner, T. E. (2002). Structure of the anxiety sensitivity index psychometrics and factor structure in a community sample. *Journal of Anxiety Disorders*, 16, 33–49. [https://doi.org/10.1016/S0887-6185\(01\)00087-1](https://doi.org/10.1016/S0887-6185(01)00087-1)
- Schneider, R. L., Arch, J. J., & Wolitzky-Taylor, K. B. (2015). The state of personalized treatment for anxiety disorders: a systematic review of treatment moderators. *Clinical Psychology Review*, 38, 39–54. <https://doi.org/10.1016/j.cpr.2015.02.004>
- Schönbrodt, F. D., & Perugini, M. (2013). At what sample size do correlations stabilize? *Journal of Research in Personality*, 47, 609–612. <https://doi.org/10.1016/j.jrp.2013.05.009>
- Seligman, M. E., & Johnston, J. C. (1973). A cognitive theory of avoidance learning. In *Contemporary approaches to conditioning and learning* (pp. xii, 321–xii, 321). Oxford, England: V. H. Winston & Sons.
- Seligman, M. E. (1972). Learned helplessness. *Annual Review of Medicine*, 23, 407–412.
- Sellström, E., & Bremberg, S. (2006). Review article: the significance of neighbourhood context to child and adolescent health and well-being: a systematic review of multilevel studies. *Scandinavian Journal of Public Health*, 34, 544–554. <https://doi.org/10.1080/14034940600551251>
- Sexton, K. A., & Dugas, M. J. (2009). Defining distinct negative beliefs about uncertainty: validating the factor structure of the intolerance of uncertainty scale. *Psychological Assessment*, 21, 176–186. <https://doi.org/10.1037/a0015827>
- Shackman, A. J., & Fox, A. S. (2016). Contributions of the central extended amygdala to fear and anxiety. *The Journal of Neuroscience*, 36, 8050–8063. <https://doi.org/10.1523/JNEUROSCI.0982-16.2016>
- Shackman, A. J., & Fox, A. S. (2018). Getting serious about variation: lessons for clinical neuroscience (a commentary on ‘The myth of optimality in clinical neuroscience’). *Trends in Cognitive Sciences*, 22, 368–369. <https://doi.org/10.1016/j.tics.2018.02.009>
- Shackman, A. J., Tromp, D. P. M., Stockbridge, M. D., Kaplan, C. M., Tillman, R. M., & Fox, A. S. (2016). Dispositional negativity: an integrative psychological and neurobiological perspective. *Psychological Bulletin*, 142, 1275–1314. <https://doi.org/10.1037/bul0000073>
- Shackman, A. J., & Wager, T. D. (2019). The emotional brain: fundamental questions and strategies for future research. *Neuroscience Letters*, 693, 68–74. <https://doi.org/10.1016/j.neulet.2018.10.012>

- Shankman, S. A., & Klein, D. N. (2003). The relation between depression and anxiety: an evaluation of the tripartite, approach-withdrawal and valence-arousal models. *Clinical Psychology Review*, 23, 605–637. [https://doi.org/10.1016/S0272-7358\(03\)00038-2](https://doi.org/10.1016/S0272-7358(03)00038-2)
- Sheynin, J., Shind, C., Radell, M., Ebanks-Williams, Y., Gilbertson, M. W., Beck, K. D., & Myers, C. E. (2017). Greater avoidance behavior in individuals with posttraumatic stress disorder symptoms. *Stress*, 20, 285–293. <https://doi.org/10.1080/10253890.2017.1309523>
- Shihata, S., McEvoy, P. M., Mullan, B. A., & Carleton, R. N. (2016). Intolerance of uncertainty in emotional disorders: what uncertainties remain? *Journal of Anxiety Disorders*. <https://doi.org/10.1016/j.janxdis.2016.05.001>
- Sidman, M. (1953). Avoidance conditioning with brief shock and no exteroceptive warning signal. *Science*, 118, 157–158.
- Sidman, M. (1962). Reduction of shock frequency as reinforcement for avoidance behavior. *Journal of the Experimental Analysis of Behavior*, 5, 247–257. <https://doi.org/10.1901/jeab.1962.5-247>
- Silverman, D. (2016). *Qualitative research*. Sage.
- Silvia, P. J., Nusbaum, E. C., Berg, C., Martin, C., & O'Connor, A. (2009). Openness to experience, plasticity, and creativity: exploring lower-order, high-order, and interactive effects. *Journal of Research in Personality*, 43, 1087–1090. <https://doi.org/10.1016/j.jrp.2009.04.015>
- Simms, L. J., Goldberg, L. R., Roberts, J. E., Watson, D., Welte, J., & Rotterman, J. H. (2011). Computerized adaptive assessment of personality disorder: introducing the cat-pd project. *Journal of Personality Assessment*, 93, 380–389. <https://doi.org/10.1080/00223891.2011.577475>
- Sims, C. R. (2018). Efficient coding explains the universal law of generalization in human perception. *Science*, 360, 652–656. <https://doi.org/10.1126/science.aaq1118>
- Sjouwerman, R., Scharfenort, R., & Lonsdorf, T. B. (2018). Individual differences in fear learning: specificity to trait-anxiety beyond other measures of negative affect, and mediation via amygdala activation. *BioRxiv*, 233528.
- Skinner, B. F. (1937). Two types of conditioned reflex: a reply to Konorski and Miller. *The Journal of General Psychology*, 16, 272–279.
- Skinner, B. F. (1963). Operant behavior. *American Psychologist*, 18, 503–515. <https://doi.org/10.1037/h0045185>
- Sleep, C. E., Hyatt, C. S., Lamkin, J., Maples-Keller, J. L., & Miller, J. D. (2018). Examining the relations among the DSM–5 alternative model of personality, the five-factor model, and externalizing and internalizing behavior. *Personality Disorders: Theory, Research, and Treatment*, 9, 379–384. <https://doi.org/10.1037/per0000240>
- Sloan, D. M., & Sandt, A. R. (2010). Depressed mood and emotional responding. *Biological Psychology*, 84, 368–374. <https://doi.org/10.1016/j.biopsycho.2010.04.004>
- Smits, D. J. M., & Boeck, P. D. (2006). From BIS/BAS to the Big Five. *European Journal of Personality*, 20, 255–270. <https://doi.org/10.1002/per.583>

- Snijders, T. A., & Berkhof, J. (2008). Diagnostic checks for multilevel models. In *Handbook of multilevel analysis* (pp. 141–175). Springer.
- Solomon, R. L., & Wynne, L. C. (1954). Traumatic avoidance learning: the principles of anxiety conservation and partial irreversibility. *Psychological Review*, *61*, 353–385.
- Spence, S. H., & Rapee, R. M. (2016). The etiology of social anxiety disorder: an evidence-based model. *Behaviour Research and Therapy*, *86*, 50–67. <https://doi.org/10.1016/j.brat.2016.06.007>
- Spielberg, J. M., Miller, G. A., Warren, S. L., Engels, A. S., Crocker, L. D., Banich, M. T., ... Heller, W. (2012). A brain network instantiating approach and avoidance motivation. *Psychophysiology*, *49*, 1200–1214. <https://doi.org/10.1111/j.1469-8986.2012.01443.x>
- Spielberg, J. M., Miller, G. A., Warren, S. L., Engels, A. S., Crocker, L. D., Sutton, B. P., & Heller, W. (2012). Trait motivation moderates neural activation associated with goal pursuit. *Cognitive, Affective, & Behavioral Neuroscience*, *12*, 308–322. <https://doi.org/10.3758/s13415-012-0088-8>
- Spielberger, C., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory, STAI*. Palo Alto, CA: Consulting Psychologists Press.
- Spira, A. P., Zvolensky, M. J., Eifert, G. H., & Feldner, M. T. (2004). Avoidance-oriented coping as a predictor of panic-related distress: a test using biological challenge. *Journal of Anxiety Disorders*, *18*, 309–323. [https://doi.org/10.1016/S0887-6185\(02\)00249-9](https://doi.org/10.1016/S0887-6185(02)00249-9)
- Sripada, R. K., Rauch, S. A. M., & Liberzon, I. (2016). Psychological mechanisms of PTSD and its treatment. *Current Psychiatry Reports*, *18*, 99. <https://doi.org/10.1007/s11920-016-0735-9>
- Stanton, K., & Watson, D. (2014). Positive and negative affective dysfunction in psychopathology. *Social and Personality Psychology Compass*, *8*, 555–567. <https://doi.org/10.1111/spc3.12132>
- Starke, J., Fineberg, N., & Stein, D. (2019). Anxiety disorders: from bench to bedside and beyond. In *Advances in Psychiatry* (pp. 33–58). Springer, Cham. https://doi.org/10.1007/978-3-319-70554-5_2
- Steenkamp, M. M., Litz, B. T., Hoge, C. W., & Marmar, C. R. (2015). Psychotherapy for military-related PTSD: a review of randomized clinical trials. *JAMA*, *314*, 489–500. <https://doi.org/10.1001/jama.2015.8370>
- Stefanis, N. C., Hanssen, M., Smirnis, N. K., Avramopoulos, D. A., Evdokimidis, I. K., Stefanis, C. N., ... Os, J. V. (2002). Evidence that three dimensions of psychosis have a distribution in the general population. *Psychological Medicine*, *32*, 347–358. <https://doi.org/10.1017/S0033291701005141>
- Stegmueller, D. (2013). How many countries for multilevel modeling? A comparison of frequentist and Bayesian approaches. *American Journal of Political Science*, *57*, 748–761. <https://doi.org/10.1111/ajps.12001>
- Stein, D. J., & Matsunaga, H. (2006). Specific phobia: a disorder of fear conditioning and extinction. *CNS Spectrums*, *11*, 248–251. <https://doi.org/10.1017/S1092852900020721>

- Studer-Luethi, B., Jaeggi, S. M., Buschkuhl, M., & Perrig, W. J. (2012). Influence of neuroticism and conscientiousness on working memory training outcome. *Personality and Individual Differences*, 53, 44–49. <https://doi.org/10.1016/j.paid.2012.02.012>
- Sturges, L. V., Goetsch, V. L., Ridley, J., & Whittal, M. (1998). Anxiety sensitivity and response to hyperventilation challenge: physiologic arousal, interoceptive acuity, and subjective distress. *Journal of Anxiety Disorders*, 12, 103–115. [https://doi.org/10.1016/S0887-6185\(98\)00004-8](https://doi.org/10.1016/S0887-6185(98)00004-8)
- Suzuki, T., Samuel, D. B., Pahlen, S., & Krueger, R. F. (2015). DSM-5 alternative personality disorder model traits as maladaptive extreme variants of the five-factor model: an item-response theory analysis. *Journal of Abnormal Psychology*, 124, 343–354. <https://doi.org/10.1037/abn0000035>
- Sylvester, P., Lilienfeld, S. O., & LaPrairie, J. L. (2011). Differences between trait fear and trait anxiety: implications for psychopathology. *Clinical Psychology Review*, 31, 122–137. <https://doi.org/10.1016/j.cpr.2010.08.004>
- Tackett, J. L., Quilty, L. C., Sellbom, M., Rector, N. A., & Bagby, R. M. (2008). Additional evidence for a quantitative hierarchical model of mood and anxiety disorders for DSM-V: the context of personality structure. *Journal of Abnormal Psychology*, 117, 812–825. <https://doi.org/10.1037/a0013795>
- Talmi, D., Seymour, B., Dayan, P., & Dolan, R. J. (2008). Human Pavlovian–instrumental transfer. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 28, 360–368. <https://doi.org/10.1523/JNEUROSCI.4028-07.2008>
- Taylor, J. M., & Yu, M. (2002). Bias and efficiency loss due to categorizing an explanatory variable. *Journal of Multivariate Analysis*, 83, 248–263.
- Taylor, S. (1996). Nature and measurement of anxiety sensitivity: reply to Lilienfeld, Turner, and Jacob (1996). *Journal of Anxiety Disorders*, 10, 425–451. [https://doi.org/10.1016/0887-6185\(96\)00021-7](https://doi.org/10.1016/0887-6185(96)00021-7)
- Tellegen, A. (1991). Personality traits: issues of definition, evidence, and assessment. In *Thinking clearly about psychology: Essays in honor of Paul E. Meehl, Vol. 1: Matters of public interest; Vol. 2: Personality and psychopathology* (pp. 10–35). Minneapolis, MN, US: University of Minnesota Press.
- Tenenbaum, J. B., & Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences*, 24, 629–640.
- Tett, R. P., & Guterman, H. A. (2000). Situation trait relevance, trait expression, and cross-situational consistency: testing a principle of trait activation. *Journal of Research in Personality*, 34, 397–423.
- Thomas, K. M., Hopwood, C. J., Donnellan, M. B., Wright, A. G. C., Sanislow, C. A., McDevitt-Murphy, M. E., ... Morey, L. C. (2014). Personality heterogeneity in PTSD: distinct temperament and interpersonal typologies. *Psychological Assessment*, 26, 23–34. <https://doi.org/10.1037/a0034318>
- Thomas, K. M., Yalch, M. M., Krueger, R. F., Wright, A. G. C., Markon, K. E., & Hopwood, C. J. (2013). The convergent structure of DSM-5 personality trait facets and five-factor model trait domains. *Assessment*, 20, 308–311. <https://doi.org/10.1177/1073191112457589>

- Thompson, B. L., & Waltz, J. (2010). Mindfulness and experiential avoidance as predictors of posttraumatic stress disorder avoidance symptom severity. *Journal of Anxiety Disorders*, 24, 409–415. <https://doi.org/10.1016/j.janxdis.2010.02.005>
- Thompson, S. C. (1981). Will it hurt less if I can control it? A complex answer to a simple question. *Psychological Bulletin*, 90, 89–101. <https://doi.org/10.1037/0033-2909.90.1.89>
- Thurstone, L. L. (1934). The vectors of mind. *Psychological Review*, 41, 1.
- Tinoco-González, D., Fullana, M. A., Torrents-Rodas, D., Bonillo, A., Vervliet, B., Blasco, M. J., ... Torrubia, R. (2015). Conditioned fear acquisition and generalization in generalized anxiety disorder. *Behavior Therapy*, 46, 627–639. <https://doi.org/10.1016/j.beth.2014.12.004>
- Torrents-Rodas, D., Fullana, M. A., Bonillo, A., Caseras, X., Andión, O., & Torrubia, R. (2013). No effect of trait anxiety on differential fear conditioning or fear generalization. *Biological Psychology*, 92, 185–190. <https://doi.org/10.1016/j.biopsycho.2012.10.006>
- Tovote, P., Esposito, M. S., Botta, P., Chaudun, F., Fadok, J. P., Markovic, M., ... Lüthi, A. (2016). Midbrain circuits for defensive behaviour. *Nature*, 534, 206–212. <https://doi.org/10.1038/nature17996>
- Tracy, J. A., Ghose, S. S., Stecher, T., McFall, R. M., & Steinmetz, J. E. (1999). Classical conditioning in a nonclinical obsessive-compulsive population. *Psychological Science*, 10, 9–13. <https://doi.org/10.1111/1467-9280.00098>
- Treadway, M. T., Bossaller, N. A., Shelton, R. C., & Zald, D. H. (2012). Effort-based decision-making in major depressive disorder: a translational model of motivational anhedonia. *Journal of Abnormal Psychology*, 121, 553–558. <https://doi.org/10.1037/a0028813>
- Treadway, M. T., Buckholtz, J. W., Schwartzman, A. N., Lambert, W. E., & Zald, D. H. (2009). Worth the ‘EEFRT’? The effort expenditure for rewards task as an objective measure of motivation and anhedonia. *PLOS ONE*, 4, e6598. <https://doi.org/10.1371/journal.pone.0006598>
- Trull, T. J., & Durrett, C. A. (2005). Categorical and dimensional models of personality disorder. *Annual Review of Clinical Psychology*, 1, 355–380. <https://doi.org/10.1146/annurev.clinpsy.1.102803.144009>
- Tzschoppe, J., Nees, F., Banaschewski, T., Barker, G. J., Büchel, C., Conrod, P. J., ... Flor, H. (2014). Aversive learning in adolescents: modulation by amygdala–prefrontal and amygdala–hippocampal connectivity and neuroticism. *Neuropsychopharmacology*, 39, 875–884. <https://doi.org/10.1038/npp.2013.287>
- Valentiner, D. P., Telch, M. J., Petrucci, D. C., & Bolte, M. C. (1996). Cognitive mechanisms in claustrophobia: an examination of Reiss and McNally’s expectancy model and Bandura’s self-efficacy theory. *Cognitive Therapy and Research*, 20, 593–612. <https://doi.org/10.1007/BF02227963>
- van Meurs, B., Wiggert, N., Wicker, I., & Lissek, S. (2014). Maladaptive behavioral consequences of conditioned fear-generalization: a pronounced, yet sparsely studied, feature of anxiety pathology. *Behaviour Research and Therapy*, 57, 29–37. <https://doi.org/10.1016/j.brat.2014.03.009>

- van 't Veer, A. E., & Giner-Sorolla, R. (2016). Pre-registration in social psychology—a discussion and suggested template. *Journal of Experimental Social Psychology*, 67, 2–12. <https://doi.org/10.1016/j.jesp.2016.03.004>
- Vanbrabant, K., Boddez, Y., Verduyn, P., Mestdagh, M., Hermans, D., & Raes, F. (2015). A new approach for modeling generalization gradients: a case for hierarchical models. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00652>
- VanElzakker, M. B., Kathryn Dahlgren, M., Caroline Davis, F., Dubois, S., & Shin, L. M. (2014). From Pavlov to PTSD: the extinction of conditioned fear in rodents, humans, and anxiety disorders. *Neurobiology of Learning and Memory*, 113, 3–18. <https://doi.org/10.1016/j.nlm.2013.11.014>
- Vasey, M. W., Harbaugh, C. N., Fisher, L. B., Heath, J. H., Hayes, A. F., & Bijttebier, P. (2014). Temperament synergies in risk for and protection against depressive symptoms: a prospective replication of a three-way interaction. *Journal of Research in Personality*, 53, 134–147. <https://doi.org/10.1016/j.jrp.2014.09.005>
- Vasey, M. W., Harbaugh, C. N., Lonigan, C. J., Phillips, B. M., Hankin, B. L., Willem, L., & Bijttebier, P. (2013). Dimensions of temperament and depressive symptoms: replicating a three-way interaction. *Journal of Research in Personality*, 47, 908–921. <https://doi.org/10.1016/j.jrp.2013.09.001>
- Vasey, M. W., & Thayer, J. F. (1987). The continuing problem of false positives in repeated measures ANOVA in psychophysiology: a multivariate solution. *Psychophysiology*, 24, 479–486.
- Veith, A. C., Russell, T. D., & King, A. R. (2017). PID-5 trait mediation of childhood maltreatment effects. *Personality and Individual Differences*, 104, 58–63. <https://doi.org/10.1016/j.paid.2016.07.024>
- Vervliet, B., & Geens, M. (2014). Fear generalization in humans: impact of feature learning on conditioning and extinction. *Neurobiology of Learning and Memory*, 113, 143–148. <https://doi.org/10.1016/j.nlm.2013.10.002>
- Vervliet, B., Vansteenwegen, D., & Eelen, P. (2004). Generalization of extinguished skin conductance responding in human fear conditioning. *Learning & Memory*, 11, 555–558. <https://doi.org/10.1101/lm.77404>
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's statement on p-values: context, process, and purpose. *The American Statistician*, 70, 129–133.
- Watson, D. (2005). Rethinking the mood and anxiety disorders: a quantitative hierarchical model for DSM-V. *Journal of Abnormal Psychology*, 114, 522–536. <https://doi.org/10.1037/0021-843X.114.4.522>
- Watson, D., & Clark, L. A. (1992). On traits and temperament: general and specific factors of emotional experience and their relation to the five-factor model. *Journal of Personality*, 60, 441–476. <https://doi.org/10.1111/j.1467-6494.1992.tb00980.x>
- Watson, D., Clark, L. A., & Harkness, A. R. (1994). Structures of personality and their relevance to psychopathology, 14.
- Watson, D., & Naragon-Gainey, K. (2014). Personality, emotions, and the emotional disorders. *Clinical Psychological Science*, 2, 422–442. <https://doi.org/10.1177/2167702614536162>

- Watson, D., Stanton, K., & Clark, L. A. (2017). Self-report indicators of negative valence constructs within the research domain criteria (RDoC): a critical review. *Journal of Affective Disorders*, 216, 58–69. <https://doi.org/10.1016/j.jad.2016.09.065>
- Watson, D., Stanton, K., Khoo, S., Ellickson-Larew, S., & Stasik-O'Brien, S. M. (2019). Extraversion and psychopathology: a multilevel hierarchical review. *Journal of Research in Personality*, 81, 1–10. <https://doi.org/10.1016/j.jrp.2019.04.009>
- Watson, D., Stasik, S. M., Ellickson-Larew, S., & Stanton, K. (2015). Extraversion and psychopathology: a facet-level analysis. *Journal of Abnormal Psychology*, 124, 432–446. <https://doi.org/10.1037/abn0000051>
- Watson, J. B., & Rayner, R. (1920). Conditioned emotional reactions. *Journal of Experimental Psychology*, 3, 1–14. <https://doi.org/10.1037/h0069608>
- Waugh, M. H., Hopwood, C. J., Krueger, R. F., Morey, L. C., Pincus, A. L., & Wright, A. G. C. (2017). Psychological assessment with the DSM–5 alternative model for personality disorders: tradition and innovation. *Professional Psychology: Research and Practice*, 48, 79–89. <https://doi.org/10.1037/pro0000071>
- Webster, M. M., & Ward, A. J. W. (2011). Personality and social context. *Biological Reviews*, 86, 759–773. <https://doi.org/10.1111/j.1469-185X.2010.00169.x>
- Weisberg, Y. J., DeYoung, C. G., & Hirsh, J. B. (2011). Gender differences in personality across the ten aspects of the Big Five. *Frontiers in Psychology*, 2. <https://doi.org/10.3389/fpsyg.2011.00178>
- Weston, C., Gandell, T., Beauchamp, J., McAlpine, L., Wiseman, C., & Beauchamp, C. (2001). Analyzing interview data: the development and evolution of a coding system. *Qualitative Sociology*, 24, 381–400. <https://doi.org/10.1023/A:1010690908200>
- Widiger, T. A. (2009). Neuroticism. In *Handbook of individual differences in social behavior* (pp. 129–146). New York, NY, US: The Guilford Press.
- Widiger, T. A., Sellbom, M., Chmielewski, M., Clark, L. A., DeYoung, C. G., Kotov, R., ... Wright, A. G. C. (2019). Personality in a hierarchical model of psychopathology. *Clinical Psychological Science*, 7, 77–92. <https://doi.org/10.1177/2167702618797105>
- Widiger, T. A., & Shea, T. (1991). Differentiation of Axis I and Axis II disorders. *Journal of Abnormal Psychology*, 100, 399–406. <https://doi.org/10.1037/0021-843X.100.3.399>
- Wilkins, K. C., Lang, A. J., & Norman, S. B. (2011). Synthesis of the psychometric properties of the PTSD checklist (PCL) military, civilian, and specific versions. *Depression and Anxiety*, 28, 596–606. <https://doi.org/10.1002/da.20837>
- Wilson, K. A., & Hayward, C. (2006). Unique contributions of anxiety sensitivity to avoidance: a prospective study in adolescents. *Behaviour Research and Therapy*, 44, 601–609. <https://doi.org/10.1016/j.brat.2005.04.005>
- Witt, L. A. (2002). The interactive effects of extraversion and conscientiousness on performance. *Journal of Management*, 28, 835–851.
- Witt, L. A., Burke, L. A., Barrick, M. R., & Mount, M. K. (2002). The interactive effects of conscientiousness and agreeableness on job performance. *Journal of Applied Psychology*, 87, 164.
- Wolpe, J., & Lang, P. J. (1964). A fear survey schedule for use in behaviour therapy. *Behaviour Research and Therapy*, 2, 27–30.

- Wong, A. H. K., & Lovibond, P. F. (2017). Rule-based generalisation in single-cue and differential fear conditioning in humans. *Biological Psychology*.
<https://doi.org/10.1016/j.biopsycho.2017.08.056>
- Wright, A. G. C., & Simms, L. J. (2014). On the structure of personality disorder traits: conjoint analyses of the CAT-PD, PID-5, and NEO-PI-3 trait models. *Personality Disorders: Theory, Research, and Treatment*, 5, 43–54.
<https://doi.org/10.1037/per0000037>
- Wright, A. G. C., Thomas, K. M., Hopwood, C. J., Markon, K. E., Pincus, A. L., & Krueger, R. F. (2012). The hierarchical structure of DSM-5 pathological personality traits. *Journal of Abnormal Psychology*, 121, 951.
- Wu, K. D., Clark, L. A., & Watson, D. (2006). Relations between obsessive–compulsive disorder and personality: beyond Axis I–Axis II comorbidity. *Journal of Anxiety Disorders*, 20, 695–717. <https://doi.org/10.1016/j.janxdis.2005.11.001>
- Xia, W., Dymond, S., Lloyd, K., & Vervliet, B. (2017). Partial reinforcement of avoidance and resistance to extinction in humans. *Behaviour Research and Therapy*. <https://doi.org/10.1016/j.brat.2017.04.002>
- Xia, Y., Gurkina, A., & Bach, D. R. (2019). Pavlovian-to-instrumental transfer after human threat conditioning. *Learning & Memory*, 26, 167–175.
<https://doi.org/10.1101/lm.049338.119>
- Yancey, J. R., Venables, N. C., & Patrick, C. J. (2016). Psychoneurometric operationalization of threat sensitivity: relations with clinical symptom and physiological response criteria. *Psychophysiology*, 53, 393–405.
<https://doi.org/10.1111/psyp.12512>
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology*, 18, 459–482.
- Zedeck, S. (1971). Problems with the use of “moderator” variables. *Psychological Bulletin*, 76, 295–310. <https://doi.org/10.1037/h0031543>
- Zinbarg, R. E., Barlow, D. H., & Brown, T. A. (1997). Hierarchical structure and general factor saturation of the anxiety sensitivity index: evidence and implications. *Psychological Assessment*, 9, 277–284. <https://doi.org/10.1037/1040-3590.9.3.277>
- Zoellner, L. A., Echiverri, A., & Craske, M. G. (2000). Processing of phobic stimuli and its relationship to outcome. *Behaviour Research and Therapy*, 38, 921–931.
[https://doi.org/10.1016/S0005-7967\(99\)00129-1](https://doi.org/10.1016/S0005-7967(99)00129-1)
- Zoellner, L. A., Pruitt, L. D., Farach, F. J., & Jun, J. J. (2014). Understanding heterogeneity in PTSD: fear, dysphoria, and distress. *Depression and Anxiety*, 31, 97–106. <https://doi.org/10.1002/da.22133>
- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67, 301–320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>
- Zuj, D. V., & Norrholm, S. D. (2019). The clinical applications and practical relevance of human conditioning paradigms for posttraumatic stress disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 88, 339–351.
<https://doi.org/10.1016/j.pnpbp.2018.08.014>

Appendix A

Table A. Summary of predictors and models tested for Aim 1.

DV: Startle and Risk Ratings											
	<u>N (W,V)</u>	<u>C (In, Or)</u>	<u>E (As, En)</u>	<u>Multi-T (norm)</u>	<u>Multi-A</u>	<u>NA</u>	<u>DE</u>	<u>DI</u>	<u>Multi-T (path)</u>	<u>DPB</u>	<u>Multi-T (PB)</u>
1	N			N,C,E	W,V,In,Or,As,En	NA			NA, DE, DI	DPB	DPB, IPB, OPB, EPB,APB
2	W,V										
3	W										
4	V										
DV: Avoidance											
	<u>N (W,V)</u>	<u>C (In, Or)</u>	<u>E (As, En)</u>	<u>Multi-T (norm)</u>	<u>Multi-A</u>	<u>NA</u>	<u>DE</u>	<u>DI</u>	<u>Multi-T (path)</u>	<u>DPB</u>	<u>Multi-T (PB)</u>
1	N	C	E	N,C,E	W,V,In,Or,As,En	NA	DE	DI	NA, DE, DI	DPB	DPB, IPB, OPB, EPB,APB
2	W,V	In,Or	As,En								
3	W	In	As								
4	V	Or	En								

Note: Each column corresponds to a conceptually-linked set of analyses (e.g., those related to Neuroticism and its aspects), and each row corresponds to a series of 3 models that allow testing of main effects (1 model), Trait x Stimulus interactions (1 model), and Trait x Stimulus² interactions (1 model). Row numbers are arbitrary and are only included for organizational purposes. Cell content indicates which trait(s) are being tested, and models with multiple traits separated by commas are modeled simultaneously. Aspects for corresponding BFAS trait are listed in parentheses. Composite traits refer to those derived from a combination of PID-5 and BFAS scales, and are indicated with a suffix of "PB". Startle and Risk Rating models were constructed separately for each DV. As = Assertiveness, APB = Assertiveness PID-BFAS, C = Conscientiousness, DE = Detachment, DI = Disinhibition, DPB = Distress PID-BFAS, En = Enthusiasm, EPB = Enthusiasm PID-BFAS, E = Extraversion, In = Industriousness, IPB = Industriousness PID-BFAS, OPB = Orderliness PID-BFAS, Or = Orderliness; Multi-A = Multi-aspect; Multi-T (path) = Multi-trait (pathological); Multi-T (PB) = Multi-trait (PID-BFAS composite); Multi-T (norm) = Multi-trait (normative); N = Neuroticism, W = Withdrawal, V = Volatility.

Appendix B

Table B1. Descriptive statistics for all BFAS scales.

Variable	Mean	SD	Min	Median	Max
Neuroticism	2.7	0.72	1.15	2.67	4.6
Withdrawal (BFAS)	2.82	0.75	1.1	2.8	4.6
Volatility	2.57	0.82	1	2.6	4.9
Agreeableness	3.99	0.56	1.5	4.05	5
Compassion	4.05	0.67	1.5	4.15	5
Politeness	3.93	0.59	1.5	4	5
Conscientiousness	3.49	0.57	1.35	3.5	4.75
Industriousness	3.46	0.67	1.2	3.5	5
Orderliness	3.53	0.64	1.5	3.6	4.8
Extraversion	3.51	0.64	1.7	3.52	4.8
Assertiveness	3.38	0.71	1.6	3.4	5
Enthusiasm	3.65	0.77	1.1	3.7	5
Openness/Intellect	3.68	0.51	1.7	3.7	4.8
Openness	3.7	0.61	1.8	3.7	5
Intellect	3.66	0.64	1.4	3.7	5

Note: Subordinate aspects follow the corresponding trait. Withdrawal is marked with "BFAS" to differentiate it from a scale on another questionnaire with the same label. BFAS = Big Five Aspect Scale.

Table B2. Descriptive statistics for all PID-5 scales.

Variable	Mean	SD	Min	Median	Max
<u>Facets</u>					
Anhedonia	0.83	0.59	0	0.75	2.62
Anxiousness	1.33	0.73	0	1.22	3
Attention Seeking	1.11	0.69	0	1	3
Callousness	0.38	0.44	0	0.21	2.57
Deceitfulness	0.79	0.56	0	0.75	2.4
Depressivity	0.51	0.57	0	0.29	2.57
Distractibility	0.99	0.66	0	0.89	2.78
Eccentricity	0.85	0.78	0	0.62	3
Emotional Lability	1.05	0.76	0	0.86	3
Grandiosity	0.65	0.56	0	0.5	2.5
Hostility	0.82	0.55	0	0.7	2.9
Impulsivity	0.8	0.64	0	0.67	3
Intimacy Avoidance	0.61	0.64	0	0.5	3
Irresponsibility	0.46	0.41	0	0.43	2.14
Manipulativeness	0.94	0.71	0	1	3
Perceptual Dysregulation	0.55	0.48	0	0.42	2.58
Perseveration	0.85	0.55	0	0.89	2.33
Restricted Affectivity	0.95	0.68	0	0.86	3
Rigid Perfectionism	1.01	0.68	0	1	2.8
Risk Taking	1.39	0.52	0.21	1.32	2.93
Separation Insecurity	0.84	0.7	0	0.71	3
Submissiveness	1.36	0.66	0	1.25	3
Suspiciousness	0.92	0.51	0	0.86	2.29
Unusual Beliefs	0.52	0.53	0	0.38	2.38
Withdrawal (PID-5)	0.79	0.61	0	0.7	3
<u>Traits</u>					
Antagonism	0.74	0.42	0.06	0.68	2.17
Detachment	0.74	0.44	0.04	0.67	2.17
Disinhibition	1.01	0.37	0.28	0.96	2.33
Psychoticism	0.66	0.54	0	0.56	2.42
Negative Affectivity	0.9	0.42	0.12	0.86	2.11

Note: Facets/traits listed in alphabetical order instead of organized by facet/trait relationship, due to many facets loading strongly on multiple traits. Withdrawal is marked with "PID-5" to differentiate it from a scale on another questionnaire with the same label. PID-5 = Personality Inventory for DSM-5.

Table B3. Descriptive statistics for all narrowband trait and state scales.

Variable	Mean	SD	Min	Median	Max
<u>Trait</u>					
TF-44	62.08	21.78	0	62	122
IUSF	31.07	8.99	12	31	59
ASI	22.4	12.15	0	20	59
STAI-T	41.95	9.94	24	41	70
<u>State</u>					
BDI-II	8.88	7.6	0	6	39
PSS	16.01	6.91	2	15	33
STAI-S	35.13	8.26	23	34	66

Note: ASI = Anxiety Sensitivity Index; BDI-II = Beck Depression Inventory-II; IUSF = Intolerance of uncertainty - Short Form; PSS = Perceived Stress Scale; TF-44 = Trait Fear - 44 item; STAI-S = State Trait Anxiety Inventory - State version; STAI-T = State Trait Anxiety Inventory - Trait version.